



University
of Glasgow

Elizalde, Aldo (2016) *Pre-colonial institutions and long-run development in Latin America*. PhD thesis.

<http://theses.gla.ac.uk/7561/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Pre-Colonial Institutions and Long-Run Development in Latin America

Aldo Elizalde

Submitted in fulfilment of the requirements for
the Degree of Doctor of Philosophy

Adam Smith Business School

University of Glasgow

September 2016

Abstract

The present doctoral thesis studies the association between pre-colonial institutions and long-run development in Latin America. The thesis is organised as follows:

Chapter 1 places the motivation of the thesis by underlying relevant contributions in the literature on long-run development. I then set out the main objective of the thesis, followed by a brief outline of it.

In Chapter 2, I study the effects of pre-colonial institutions on present-day socioeconomic outcomes for Latin America. The main thesis of this chapter is that more advanced pre-colonial institutions relate to better socioeconomic outcomes today - principally, but not only, through their effects on the Amerindian population. I test such hypothesis with a dataset of 324 sub-national administrative units covering all mainland Latin American countries. The extensive range of controls covers factors such as climate, location, natural resources, colonial activities and pre-colonial characteristics - plus country fixed effects. Results strongly support the main thesis.

In Chapter 3, I further analyse the association between pre-colonial institutions and present-day economic development in Latin America by using the historical ethnic homelands as my main unit of analysis. The main hypothesis is that ethnic homelands inhabited by more advanced ethnic groups -as measured by their levels of institutional complexity- relate to better economic development today. To track these long-run effects, I construct a new dataset by digitising historiographical maps allowing me to pinpoint the geospatial location of ethnic homelands as of the XVI century. As a result, 375 ethnic homelands are created. I then capture the levels of economic development at the ethnic homeland level by making use of alternative economic measures –satellite light density data. After controlling for country-specific characteristics and applying a large battery of geographical, locational, and historical factors, I found that the effects of pre-colonial institutions relate to a higher light density –as a proxy for economic activity- in ethnic homelands where more advanced ethnic groups lived.

In Chapter 4, I explore a mechanism linking the persistence of pre-colonial institutions in Latin America over the long-run: Colonial and post-colonial strategies along with the

ethnic political capacity worked in tandem allowing larger Amerindian groups to "support" the new political systems in ways that would benefit their respective ethnic groups as well as the population at large. This mechanism may have allowed the effects of pre-colonial institutions to influence socioeconomic development outcomes up to today. To shed lights on this mechanism, I combine the index of pre-colonial institutions prepared for the second chapter of the present thesis with individual-level survey data on people's attitudes. By controlling for key observable and unobservable country-specific characteristics, the main empirical results show that areas with a history of more advanced pre-colonial institutions increase the probability of individuals supporting present-day political institutions.

Finally, in Chapter 5, I summarise the main findings of the thesis, and emphasise the key weaknesses of the study as well as potential avenues for future research.

Contents

Abstract	ii
Contents	iii
Table of Contents	iv
List of Tables	v
List of Figures	vi
Acknowledgements	vii
Declaration	viii
1 Introduction	1
1.1 Motivation and main objective of the thesis	1
1.2 Outline of the thesis	6
2 Pre-colonial Institutions and Socioeconomic Development: The Case of Latin America	8
2.1 Introduction	8
2.2 Historical overview and main hypothesis	11
2.3 Data and methodology	14
2.4 Empirical analysis	23
2.5 Concluding remarks	36
3 Pre-Colonial Institutions, Ethnic Homelands, and Economic Development in Latin America	38
3.1 Introduction	38
3.2 Related literature	41
3.3 Historical context	42
3.4 Data and Methodology	44
3.5 Empirical strategy and main results	55

3.6	Conclusions	63
4	On the Mechanisms of Long-Run Effects of Pre-Colonial Institutions in Latin America	65
4.1	Introduction	65
4.2	Related literature	67
4.3	Historical context	69
4.4	Data and Methodology	73
4.5	Empirical Results	79
4.6	Conclusions	85
5	Conclusions	86
5.1	Summary of key findings of the thesis	86
5.2	Challenges and potential avenues for future research	87
	Bibliography	89
A	Appendix –Chapter 2–	94
B	Appendix –Chapter 3–	102
C	Appendix –Chapter 4–	105

List of Tables

2.1	Latin American countries and their Amerindian populations	20
2.2	Descriptive Statistics	21
2.3	Matrix of correlations	22
2.4	Baseline results	24
2.5	Baseline results with 8 different dependent variables	27
2.6	Controlling for colonial activities	30
2.7	Controlling for pre-colonial characteristics: population density	32
2.8	Controlling for pre-colonial characteristics: socioeconomic factors	34
2.9	Contrasting rural and urban areas	36
3.1	Summary statistics	56
3.2	Baseline results	58
3.3	Pre-Colonial Development and Colonial Activities	60
3.4	Pre-Colonial Institutions (dummy approach)	62
3.5	Robustness checks	64
4.1	Summary statistics	78
4.2	Pre-Colonial Institutions and Population Supporting the State	80
4.3	Pre-Colonial Institutions and Contemporary Political Institutions	82
4.4	Population groups	83
4.5	Robustness checks (Logit vs Probit)	84
A.1	Ethnic groups	95
A.2	Definitions and sources of variables (baseline analysis)	98
A.3	Description of additional pre-colonial characteristics	100
A.4	Contrasting rural and urban areas	101
B.1	Definitions and sources of variables	103
C.1	Definitions and sources of variables	106

List of Figures

2.1	Pre-Colonial Institutions in Latin America	17
3.1	Traditional Ethnic Homelands in Latin America	46
3.2	Pre-Colonial Institutions within Ethnic Homelands in Latin America . . .	49
3.3	Light Density at Night within Ethnic Homelands in Latin America	51
3.4	Intersection between Ethnic Homelands and Subnational States (Colombia)	54
3.5	Categories of Pre-Colonial Institutions and Light Density at Night	61

Acknowledgements

I dedicate my thesis to my partner, Ruth, and lovely children, Mhairi, Callum, and Niamh, whose support, motivation and love made it work.

My gratitude also goes to my mum, for her invaluable maternal support, as well as to my sister, Marycarmen, and brothers, Juan and Jorge, for their faithful encouragement.

I would also like to thank my main supervisor, Dr Luis Angeles, for his continuing guidance as well as outstanding supervisory role.

Special thanks to my examiners, Dr. Daniel Oto- Peralías and Dr Céline Azémar, for their time and helpful comments on my work.

I am also thankful to my best friend in my PhD studies, Timothy Birabi, for his productive and intensive debate on my work.

Finally, I gratefully acknowledge financial support from the National Council for Science and Technology of Mexico (CONACyT), in order to undertake my PhD studies at the University of Glasgow.

Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Printed name: Aldo Elizalde

Signature:

Chapter 1

Introduction

1.1 Motivation and main objective of the thesis

"So catastrophic and irrevocable were the changes that is tempting to think that almost nothing survived from the past. This is wrong: landscape and people remain, though greatly altered. And they have lessons to heed, both about the earth on which we all live, and about the mental frames we bring to it".

(Mann 2005, p. 315)

Understanding of how developing countries got to where they are now is becoming a devoted line of research within the economic literature these days. Knowing this rightly allows us to see not only *where* we want head in the future but also *how* we want to move forward. Yet until recently, scholars interested in this field of study have examined the long-run of wealth and poverty of the Developing World from a particular historical perspective: the role of colonial institutions.

The emphasis of such colonial approach is on the transmission of forces coming from the Western civilisation via the phenomenon of colonisation. The main arguments of this approach are based on the idea that developing countries are nowadays rich or poor depending on what Europeans brought and the different ways by which the new settlements took place. For example, amongst other factors, Europeans brought their own language and legal systems with them and, as a result, a rather different set of institutional framework in the Developing World was put in place. These factors are then ideal sources of exogenous variation to explain econometrically the institutional differences affecting the paths of economic development over the long-run.

A seminal paper studying the effects of colonial institutions on contemporary development is that of Engerman and Sokoloff (1997). They studied, in a narrative way, the long-run development in the Americas by looking at the ways that colonisation took off in

front of factor endowments. They put forward the hypothesis that places with lands rich in mineral resources such as gold and silver along with advanced pre-colonial societies, developed institutions that supported forced labour activities. Similarly, colonies that were endowed with climate and land suitable for trading commodities, developed plantation activities giving rise to the use of slave labour. The new institutional arrangements in place then established an uneven socioeconomic system based on a concentration of landownership and an intensive extraction of natural resources. All this resulted in the formation of forces that gave way to social and economic inequalities that have persisted to our days –in particular in the area what we know today as Latin America.

In a later seminal paper of Acemoglu et al. (2001), a ground-breaking hypothesis was introduced placing the role of colonial institutions as one of the most relevant drivers in shaping economic development in the long-run. Acemoglu et al. (2001) argued that the type of colonial institutions in place all along the colonised World varied as a result of geographical determinants –in particular the disease environment. Using countries as their main unit of analysis across the globe, they put forward the thesis that in colonies with temperate environments the degree of European mortality rates was significantly lower than in colonies with more tropical environments such as Central America or sub-Saharan Africa. This, in turn, resulted in a higher European settlement. As such, a type of colonial rule based mainly on Western values and beliefs was then put in place, giving way to a institutional framework with robust mechanisms of checks and balances. On the other hand, colonies with more tropical environments increased the European mortality rates as a result of diseases such as malaria. Hence, European settlement did not take the same proportion. In these colonies, Europeans pursued extractive methods relying on –most of the time– the numerous native population and social organisation to extract resources in the detriment of everything else. Under this adverse context, the colonial rule lacked of incentives in setting up robust institutions similar as the ones in place in the colonies with more pro-Western environments. Acemoglu et al. (2001) used European mortality rates as an instrument for today’s institutions –specifically they used a measure of expropriation risk as a proxy for institutions on property rights– to capture the casual effect of institutions on contemporary development in a cross-section of countries at global scale.

Followed the works of Engerman and Sokoloff (1997) and Acemoglu et al. (2001), an extensive and notable volume of studies stressing the role of colonial institutions dominated our attention as to what and how we should understand the paths of economic development in the long-run. Indeed, the colonial experience changed the colonised World in many ways. French, English, Spanish or Portuguese colonisers put in place values and beliefs resembling those from their country of origin, for which case *invariably* the life and space of the native population of the Developing World got to be transformed ever since. While

this is true as documented in the literature, we must not forget that these colonies were the homeland of a large variety of societies with their own special set of social characteristics. Thus, as the Western values and beliefs shaped the paths of economic development in the colonised World over the long-run, this should also be the case when considering the social characteristics of the societies that lived before colonisation –and of course with its own type of degree and settings.

Certainly, scholars have not disregarded entirely what was before colonisation in searching for the ultimate drivers of development over the long-run. Indeed, some studies do take into account this part of history. Yet somehow this analysis focuses principally on making the case for the persistence of the colonial experience. A good example of this line of research is that of Acemoglu et al. (2002). They put forward the hypothesis that the richest pre-colonial societies are today amongst the poorest areas in the World. Using a cross-section of countries, they capture the development of societies prior colonisation by making use of a measure of population density in 1500. The historical narrative as to how this *reversal of fortune* came to happen is simple to grasp. Typically, advanced pre-colonial societies such as the Incas or Aztecs in the Americas had cities highly urbanised. After the Conquest, such element of urbanisation was essential for the colonisers in the extraction of natural resources. As noted before, this type of colonial rule set up extractive strategies over time. On the other hand, in areas considered as less advanced or poorer prior colonisation -as reflected by population density in 1500-, the colonial rule somewhat reversed the institutional settings already in place there resulting in countries where pro-development strategies were the rule to follow. The divergence in development started to take off by the 19th century when the opportunities for countries to get industrialised were only possible in those areas where the colonial rule gave way to institutions based on Western values and beliefs. The key rationale of this study is *not* to capture a direct effect of pre-colonial factors on present-day development but instead via the colonial experience, placing once again at centre of the discussion the superiority of colonial institutions.

It is perhaps understandable, however, why the studies addressing the development in the long-run have not yet factored in the direct effect of pre-colonial factors on present-day development. The paucity of data as far as 500 years ago in most of the colonised World is quite usual. While exceptional efforts in constructing adequate historical measures of development have been done, scholars still need to deal with key empirical challenges in light of the almost unavoidable scenario of omitted-variable bias; in which case country-level studies are the most affected. As such, recent studies have shifted the analysis to a finer empirical inspection defined at the subnational level. Yet again the focus of these studies is essentially on the role of colonial institutions.

A key paper studying the long-run development at subnational level is that of Bruhn

and Gallego (2012). They use a cross-section of 345 subnational states in the Americas to study the association between the main economic activities in place during the early colonial period and present-day development. Supported by the main findings of Engerman and Sokoloff (1997), they argue that the colonial economic activities based on exploited labour such as mining, plantation, sugar, etc. –what they called as "bad" activities– is negatively correlated with contemporary development. Moreover, and relying on a novel measure of pre-colonial population density –defined at subnational level–, they find that regions where other colonial economic activities were undertaken like cattle or farming, are adversed for development over the long-run *only* if these regions had large native population to be exploited. Additionally, and as noted in Acemoglu et al. (2002), they also capture a strong negative association between their measure of pre-colonial population density and present-day development -as reflected *principally* by income per capita.

In effect, Bruhn and Gallego’s findings echo the line of research placing the role of colonial institutions as a key driver for long-run development in the Developing World. Futhermore, it is noteworthy to highlight that these type of studies have enabled sholars to control for a sizable effect of revelant unobservable country-specific characteristics affecting the variation of development by means of country-fixed effects. For example, national factors other than those specifically controlled for may be behind the statistical relationship between colonial institutions and long-run development. Subnational level studies -and in specific the application of country-fixed effects- seem to address this empirical issue.

Turning our attention to the main hypothesis that the present doctoral thesis is putting forward. While the process of colonialism modified the reality and fortune of the societies that Europeans encountered at their arrival, some pre-colonial factors survived and, thus, interacted with the colonial rule –as rightly pointed out already in the literature. What remained then, I believe, may have carried the persistence of pre-colonial culture and institutions all along the colonial and post-colonial periods, and thus have affected the long-run development in the Developing World as well.

And indeed, the above assertion has been the line of study of a novel strand in the literature analysing the direct effects of pre-colonial factors on long-run development of developing countries. Yet all these studies have directed their attention to the African case. A starting point of this new set of studies is the work of Gennaioli and Rainer (2007), who uncover a strong positive association between pre-colonial institutions and various measures of socioeconomic outcomes in a cross-section of countries mainly from sub-Saharan Africa. They measure institutions using a categorical variable that classifies ethnic groups according to their level of political complexity drawn from the pioneering anthropological data of George Peter Murdock’s Ethnographic Atlas (Murdock 1967). Nevertheless, as noted before, country level studies like this one may be more questionable

due to the large unobservable country-specific characteristics left unquantified, and as such they may be possibly behind this relationship.

In a later study, Michalopoulos and Papaioannou (2013) advanced the work of Gennaioli and Rainer (2007) by using ethnic groups as their main unit of analysis in order to explore the relationship between pre-colonial institutions and long-run development in the Africa context. As in Gennaioli and Rainer (2007), Michalopoulos and Papaioannou (2013) used the same anthropological variable of Murdock’s Ethnographic Atlas in order to measure the institutions from the pre-colonial period. In addition to confirming the main hypothesis that was introduced by Gennaioli and Rainer (2007), the unit of analysis -ethnic groups- used by Michalopoulos and Papaioannou (2013) allowed them to apply country-fixed effects as a relevant empirical strategy to control for unobservable national factors.

As correctly emphasised by Gennaioli and Rainer (2007) and Michalopoulos and Papaioannou (2013), somewhat a few more studies have stressed these days the relevance of other pre-colonial factors in development over the long-run –principally in the Africa context. For example, Nunn (2008) uncovered a strong negative association between slave trade undertaken during the pre-colonial period and contemporary development in Africa. Yet what it is noteworthy to remark at this point is that all these studies have allowed us to rethink our understanding about the way that societies got to where they are now from a different historical perspective: the role of pre-colonial characteristics. Recognising, and most important, knowing the extent to which this other layer of social characteristics has affected the development over the long-run *must* be weighed up together with the colonial aspects. And this is what the present doctoral thesis aims.

Thus, the main objective of the present doctoral thesis is to study, *specifically*, the effects of pre-colonial institutions on long-run development in the context of **Latin America**¹. Quite rightly, in the Latin American case different realities and historical events arise as compared to the African setting. Colonialism was not only much longer lasting than in other regions of the world, about three centuries for most Latin American nations, but it was also accompanied by a massive transformation of the ethnic structure of the population. While correct, this should not lead us to conclude that pre-colonial culture was simply wiped out following the European arrival and had no chance of influencing the present. Western culture may well be the dominant element in present-day Latin America, as best exemplified by the almost universal adoption of the Christian religion

¹The thesis will not be covering countries in North America. This is indeed a research choice. The countries I will cover were all colonized by Spain, with the exception of Brazil which was colonized by Portugal. I believe this results in a similar colonial experience and a significant degree of cultural proximity across the countries analysed. Including Canada or the United States would render the sample substantially more diverse, and results could be driven by these two countries.

and the languages from Portugal and Spain, but in numerous cases this dominance contains important elements of the native culture. Pre-colonial rites, such as offerings to the Mother-Earth, are still common among large sections of the Latin American population –and do not stop those who perform them from attending the Sunday Mass. And pre-colonial languages such as *quechua* or *nahuatl* (the lingua francas of the former Inca and Aztec empires) can be easily heard in the food markets of cities and towns across the Andes and central Mexico. Elements of pre-colonial culture have survived five centuries of colonial and post-colonial regimes, and the present thesis discusses how this took place and why it could matter for the analysis of current socioeconomic outcomes.

1.2 Outline of the thesis

The present doctoral thesis studies the relationship between pre-colonial institutions and long-run development in Latin America. In Chapter 2 and 3, I will explore the direct effects of pre-colonial institutions on present-day development in a cross-section of subnational units in Latin America. While Chapter 2 uses subnational states as the main unit of analysis, Chapter 3 uses ethnic groups –both units though are defined below the national level. In Chapter 4, I will explore a mechanism linking the persistence of pre-colonial institutions over the long-run. Finally, Chapter 5 concludes outlining weaknesses and potential areas for future research.

As noted above, subnational level studies are currently used as a relevant way to reduce identification problems arising from historical shocks, cultural adaptations, or colonial and post-colonial institutions by means of country-fixed effects. Motivated by this type of method, in Chapter 2 I use as my main unit of analysis the largest administrative divisions of each country in Latin America below the national level called *states*. Thus, the chapter relies on a new constructed dataset by making use of statistics from 324 subnational administrative units. To measure the effects of pre-colonial institutions in each unit of observation, a novel index is then constructed with the information of the population shares of the different Amerindian groups at the state level as well as data on the political complexity of ethnic groups drawn from Murdock’s Ethnographic Atlas. The historical narrative essentially makes the case for the persistency of pre-colonial institutions. Specifically, it is argued that pre-colonial institutions survived to our days thanks to the existence of largely self-governed Amerindian communities in rural Latin America. Overall, the main result shows that more advanced pre-colonial institutions relate to better socioeconomic outcomes today - principally, but not only, through their effects on the Amerindian population.

In Chapter 3, I further analyse the association between pre-colonial institutions and

present-day economic development in Latin America from an innovative empirical fashion. I use the historical ethnic homelands as my main unit of analysis rather than subnational states. While this method is widely used in the African context, for the Latin American case there is no formal studies addressing such relationship through the effects of the historical ethnic homelands. Certainly, the major difficulty for this type of studies is the paucity of the required data. While data on the geospatial location of ethnic groups is indeed available, all the opened sources construct their own datasets on the basis of the modern location of ethnic groups around the world –including those from the Latin American countries. This, however, may make less tenable a unit of analysis that is meant to have diffused an effect 500 years ago. Knowing the historical location of the ethnic groups before colonisation may be far too more revealing when it comes to uncover an historical effect within a specific geographical area. By using Geographical Information System, in Chapter 3, I thus construct a new dataset by digitising historiographical maps allowing me to pinpoint the geospatial location of ethnic homelands as of the XVI century. The method, indeed, comes with its own methodological challenges and differences as the one from Chapter 2, yet it is very striking to see that the persistency of pre-colonial institutions is again uncovered. The Chapter shows that ethnic homelands inhabited by more advanced ethnic groups -as measured by their levels of institutional complexity- relate to better economic development today -as reflected by satellite light density data, which proxy for economic activity at the required level of analysis.

The persistency of pre-colonial institutions is studied empirically in Chapter 4 using a mechanism related to the ethnic political capacity and the main strategy of dominance implemented by the colonial and post-colonial systems. The historical narrative underlines the variation of political capacity of ethnic groups. Typically, larger ethnic groups were more likely to have endured the new systems through their highly advanced political capacity inherited from their native institutional frameworks. Additionally, in light of the overwhelming geographical conditions, colonisers set up an indirect rule resulting in numerous self-governed Amerindian communities. These two forces -ethnic political capacity and colonial and post-colonial strategies- may have given rise in areas traditionally inhabited by larger groups to some degree of "support" to the new political systems, triggering as a result better social and economic benefits in both the Amerindian communities and the rest of the population. To shed lights on this mechanism, I combine the index of pre-colonial institutions constructed in Chapter 2 with individual-level survey data on people's attitudes. The main empirical results show that areas with a history of more advanced pre-colonial institutions increase the probability of individuals supporting present-day political institutions. This degree of "support" may offer some evidence of such mechanism linking the persistency of pre-colonial institutions over the long-run.

Chapter 2

Pre-colonial Institutions and Socioeconomic Development: The Case of Latin America

Abstract

I study the effects of pre-colonial institutions on present-day socioeconomic outcomes for Latin America. The main hypothesis is that more advanced pre-colonial institutions relate to better socioeconomic outcomes today - principally, but not only, through their effects on the Amerindian population. I test the hypothesis with a dataset of 324 sub-national administrative units covering all mainland Latin American countries. The extensive range of controls covers factors such as climate, location, natural resources, colonial activities and pre-colonial characteristics - plus country fixed effects. Results strongly support the main hypothesis.

2.1 Introduction

Over the last two decades, the economics literature searching for the ultimate drivers of the wealth and poverty of nations has given considerable attention to the role of institutions. Taking the theoretical discussions of Douglass North as a starting point (North 1981; North 1990), the literature has progressed mainly along empirical lines and produced an impressive array of supporting evidence: see Hall and Jones (1999) and Acemoglu et al. (2001) for two important early contributions, and Acemoglu et al. (2005) for a review of the literature.

As has been noted before, most of this literature focuses on the role of colonial institutions - the institutional package that European colonial powers put in place throughout the world between the 16th and 20th centuries. There are indeed good reasons for focusing

on the impact of colonialism on institutional development in countries outside Europe. For a start, colonialism was often a deeply disrupting process that radically modified the way societies were structured. Factors such as the content and direction of trade, the nature of taxation, and the operation of the law were always affected. Furthermore, colonialism reached the vast majority of countries we now call the Developing World (plus a few countries from today's Developed World), making it an ideal source of exogenous variation of institutional quality in cross-country comparisons.

While the importance of colonialism is well-recognized in this context, recent research has progressively uncovered a large role for pre-colonial factors as well. Most of this research has focused on Sub-Saharan Africa, a continent where colonialism arrived late (most of the African interior was largely unknown to Europeans until the 1880s), did not last very long (African decolonisation took off during the late 1950s), and where, with the exception of South Africa and its neighbours, European settlement was of very limited importance. Under these conditions, it is perhaps not surprising that pre-colonial factors transcend the colonial period and affect African societies to this day.

A good example of this line of research is Nunn (2008), who shows how the intensity of the African slave trade over the pre-colonial period is negatively related to current levels of income per capita. Nunn and Wantchekon (2011) expand on this by uncovering a relationship between the intensity of the slave trade and present-day levels of interpersonal trust. Also of relevance, Huillery (2011) argues that the attitudes of pre-colonial African states towards Europeans have an influence on current development outcomes as colonizers invested more in the areas where Africans were less hostile.

Turning to pre-colonial institutions, Gennaioli and Rainer (2007) and Michalopoulos and Papaioannou (2013) uncover a positive link between the institutional characteristics of African ethnic groups, as prevalent in the pre-colonial period, and present-day measures of socioeconomic development. Both papers measure institutions using the same categorical variable that classifies ethnic groups according to their level of political complexity (a value of 0 denotes bands and tribes, while the maximum value of 4 is for complex states). The data is taken from George Peter Murdock's *Ethnographic Atlas* (Murdock 1967), which codes more than 60 variables for 1267 ethnic groups around the world¹.

It is important to note that the *Atlas* attempts to describe ethnic groups in the absence of foreign elements, in particular colonial influence. For the Latin American case this often implies a reliance on historical sources in addition to direct anthropological evidence. For

¹The *Ethnographic Atlas* has been expanded and updated several times since its publication. The current version, used in this paper, is due to Gray (1999) and can be accessed at http://intersci.ss.uci.edu/wiki/index.php/Ethnographic_Atlas

instance, for several important ethnic groups such as the Incas, Aztecs, Mayas, Chibchas or Tarascans, the "approximate time to which the data pertain" is given as the 1500s.

Gennaioli and Rainer (2007) use African nations as their unit of analysis and uncover a positive relationship between pre-colonial institutions and measures of health, education and other public goods. The main caveat of this approach is that national factors other than those specifically controlled for, and in particular colonial and post-colonial institutions, may be behind the uncovered statistical relationship. To address this issue Michalopoulos and Papaioannou (2013) use ethnic groups as their unit of analysis, and show how pre-colonial institutions have an effect on present-day economic development even after controlling for national characteristics with the use of country fixed effects. In subsequent work, Michalopoulos and Papaioannou (2014) also show how African national institutions, that is, institutions put in place by colonial and post-colonial forces, actually exert little or no influence in areas far away from the capital city - a result that further emphasizes the importance of pre-colonial factors.

This Chapter contributes to the literature by analysing the role of pre-colonial institutions on present-day socioeconomic outcomes for Latin America. To the best of my knowledge, this is the first attempt of this kind for the Latin American case. Because controlling for colonial and post-colonial institutions by means of country fixed effects is an important requirement, I do not follow Gennaioli and Rainer (2007) in using the nation as the unit of analysis. Michalopoulos and Papaioannou (2013) use ethnic groups as their unit analysis which is possible in the Latin American case but comes with its own methodological challenges -as we will see in the next Chapter. For this Chapter, instead, I take as my unit of analysis the largest administrative divisions of each country below the national level, which I refer to as "states"². Of course, these states are ethnically diverse; but data on the share of the different ethnic groups within them is available. The data thus covers 324 states from 17 Latin American countries, and allows for the use of country-specific fixed effects. Among this cross-section of states I uncover a robust and statistically significant relationship between pre-colonial institutions and present-day measures of education, health, and economic development. The relationship remains in place after controlling for each state's geographic characteristics and a battery of alternative colonial and pre-colonial determinants of economic success.

Besides the above references, the present Chapter is also related to a rapidly developing empirical literature using data at the sub-national level to shed new light on questions of growth and development (Acemoglu and Dell 2010; Gennaioli, La Porta, Lopez De Silanes, and Shleifer 2014). Within this literature, of special relevance to us is the work of Bruhn

²The actual name given to these administrative divisions changes from country to country: provincias in Argentina, departamentos in Bolivia, regiones in Chile, estados in Mexico, and so on.

and Gallego (2012), who not only focus on Latin America but also consider the effect of different colonial economic activities, like mining or plantation agriculture. This offers a relevant set of controls for my study, as I seek to factor out pre-colonial influences from colonial ones. A second paper of relevance to my study is Maloney and Valencia (2015), who show there is persistence of economic success at the subnational level in the Americas. Since much of this persistence is due to the fact that colonial settlements were often established near pre-colonial population centres, the work of Maloney and Valencia cannot assess the role of pre-colonial institutions on economic development - precisely the point of the present chapter. Finally, the Chapter also relates to a literature stressing the deep historical roots of long-run development. Within this literature, the evidence presented by Bockstette et al. (2002) and Comin et al. (2010) suggests that factors such as early state formation and technological advancement may determine future trajectories of economic development for us as much as two or three millennia.

The rest of the Chapter is organized as follows. Section 2.2 offers some historical background in order to place the research question in its context. In particular, it makes the case for pre-colonial institutions having an influence up to our days in Latin America. Section 2.3 discusses the data and presents the empirical methodology. Section 2.4 presents the baseline results and extends into a number of additional tests and robustness checks. Section 2.5, finally, offers some concluding remarks.

2.2 Historical overview and main hypothesis

On the face of it, it is perhaps understandable that the research effort on the role of pre-colonial institutions has so far ignored the case of Latin America. Indeed, there is little doubt that colonialism cut much deeper in Latin America than in Africa -or almost any other region in the world for that matter. It is not just that colonialism lasted far longer, about three centuries for most Latin American nations and even longer for a few Caribbean ones. More important, only in the Americas and in some of the Pacific islands did the European conquest lead to a radical transformation of the ethnic structure of the population. Suffering the consequences of a new disease environment, the aboriginal population of the Americas (henceforth *Amerindians*) was decimated over the hundred years or so following first contact with Europeans. In its place, a society of European descendants, mixed-race mestizos, Amerindians and Africans, these last ones brought as slave labour, took over. Europeans moved to the Americas permanently, and their descendants constitute the economic and political elite of most Latin American nations to this day.

Under these circumstances, one would be excused to think that Latin America con-

stitutes a prime example of how colonial factors determine institutional structures to the detriment of everything else. Only in the Americas we can observe colonialism wiping out not only the established institutional order but even the very people who held that order. The rupture from the pre-colonial world was radical, the chances for pre-colonial factors to survive into the present-day appear, at first sight, very limited.

And yet, things start to look different under closer inspection. While Amerindians remained at the fringes of the economic and political power structures, they played a crucial role as the main source of labour, and therefore principal factor of production, in two of the most important sectors of the colonial economy: mining and agricultural production for the local market³. Amerindian labour was the main source of wealth for Spanish settlers in the Americas, as best summarized by the aphorism "*Sin Indios no hay Indias*" ("*Without Indians there is no Indies*") - attributed to 16th century Spanish settlers when defending the granting of rights over Indian labour against accusations by the Crown of excessive exploitation.

The extraction of this Amerindian labour relied on the use of aboriginal structures of power and organization. While the growing class of mestizos lived in towns and cities and collaborated closely with the Spanish elites, Amerindians by and large retired to their rural communities where they lived a separate cultural life from the rest of society. Spanish governors referred to this network of Indian villages, where no Europeans lived permanently, as the "*Republic of Indians*" - a name that reveals much about the degree of autonomy granted to these communities in their internal affairs. Indian villages were compelled to pay taxes and supply tribute in the form of labour for mining, public works and, during the early phase of the colony, agricultural production through a number of schemes such as the *encomienda*, *repartimiento* or *mita*. The delivery of labour and taxes was organized by local headmen and leaders, who enjoyed privileges such as the private ownership of land and exception from taxation. In this way, as James Lang put it, "*The Spanish enterprise in the New World rested on an indigenous social order*" (Lang 1975, p. 7).

After bottoming out in the early to mid-17th century, Amerindian populations started to recover all along the continent from the early 18th century onwards (Burkholder and Johnson 1998, pp.107-110). The rise of the large agricultural estate (*hacienda*) during the late colonial period and through the 19th century meant that many Amerindians left their communities to find permanent work (and often debt bondage) in them. On the other hand, the 20th century brought a number of revolutionary movements and

³Agricultural production for the export market, with crops such as sugar cane, tobacco and cotton, employed African slave labour. See Angeles (2013) for an analysis of the factors determining the flow of slaves to the Americas.

left-leaning governments which aimed at redistributing land and in so doing reversed the flow of Amerindians out of their communities (most notably the Mexican revolution of 1910 and the Bolivian revolution of 1952). These ups and downs notwithstanding, during five centuries of colonial and post-colonial history the Amerindian rural community has remained a permanent element of Latin American countries. Its existence ensured the preservation of Amerindian languages, cultural characteristics and many institutional elements up to the present day.

The above historical overview applies to most Amerindian groups throughout the Americas - only the most remote groups such as the tropical forest dwellers of the Amazon managed to remain outside the influence of Europeans well into the 20th century. Within this large universe of Amerindian cultures, large institutional differences were in place. At the time of first contact with Europeans, Amerindian groups varied greatly in terms of political structure and institutional complexity: from the multi-layered bureaucracy administering the vast Inca Empire to the numerous small chiefdoms with no political organization beyond the village level, passing through intermediate forms of political complexity such as the confederacies of villages and city-states of Mesoamerica.

The central hypothesis of the present chapter –and the doctoral thesis as a whole– is that these institutional differences, preserved throughout the colonial period in Amerindian rural communities, continue to exert an influence on socioeconomic development up to this day. As has been shown for the case of Africa, I hypothesize that higher levels of pre-colonial institutional development are associated with better outcomes in areas such as education, health, and economic well-being. Several mechanisms have been advanced to explain this link in the literature. In the context of Latin America, I would emphasize the following ones:

i) Ethnic groups with experience of large-scale political organization were in a better position to ensure the delivery of locally-produced public goods such as education and public health. They would also have local forms of legal resolution that did not involve colonial or national courts (and would for that reason be more efficient).

ii) Institutionally-advanced groups found it easier to learn new techniques and modes of production, and to integrate themselves into the colonial and post-colonial economic system. For instance, experience with markets was much more prevalent among Aztecs and Incas than among peoples of the Amazon.

iii) Institutionally-advanced groups were able to organize themselves and defend their interests, including claims to land and other resources, in front of colonial and post-colonial governments.

iv) A higher level of pre-colonial institutional development may result in more accountability of local chiefs, in particular if some forms of political organization survived

beyond the village level which would hold village leaders accountable towards their ethnic nations. This mechanism is emphasized by Gennaioli and Rainer (2007) for the case of Africa.

The above explanations have in common that they all describe a positive link between the pre-colonial institutions of an ethnic group and the subsequent socioeconomic success of that same group. In addition to them I may also hypothesize that pre-colonial institutions could benefit the non-Amerindian population as well. This is of particular importance for Latin America, where the dominant ethnic group in most regions is the mestizo - people with mixed Amerindian and European ancestry. While mestizos clearly function within the institutional setting put in place by the colonial and post-colonial state, pre-colonial institutions may exert some influence on their beliefs, cultural practices, and ability to cooperate with ethnic groups of European origin. In this way, the influence of pre-colonial institutions may be felt across all the inhabitants of any given state, and not just Amerindians.

2.3 Data and methodology

I follow Gennaioli and Rainer (2007) and Michalopoulos and Papaioannou (2013) in using Murdock's (1967) index of "*Jurisdictional Hierarchy beyond the local community level*" as our measure of pre-colonial institutional complexity. The variable takes discrete values between 0 and 4, where the value represents the number of levels of political organization above the local community. Murdock assigns a value of 0 to groups organized in bands or single-village tribes, a value of 1 to chiefdoms comprising a few villages or a single city-state, and a value of 2 for large chiefdoms with many cities or confederacies of city-states. Values 3 and 4 are reserved for states with several levels of intermediate bureaucracy between its ruler and the local community (provinces, municipalities and so on). These categories are somewhat related to the standard classification of political complexity in anthropological studies, as first formulated by Elman Service, which classifies societies into bands, tribes, chiefdoms and states (Service 1971). As discussed by Diamond (1997), the level of political complexity is closely related to technological advancement, which is needed in order to support an ever larger class of non-food producers.

For the Americas, the only pre-colonial group that achieves the maximum value of 4 in Murdock's classification is the Incas. Indeed, the Inca Empire is well-recognized as the most sophisticated political and administrative structure developed in the Western Hemisphere before the European conquest (Burkholder and Johnson 1998, p. 19)⁴. Perhaps

⁴I also assign the value of 4 to the Aymaras, a large Amerindian group which was part of the Inca

surprisingly, the Aztec Empire of central Mexico, the only pre-colonial state comparable to the Inca Empire in terms of extension and population, is only assigned a value of 2. This, arguably, is due to the Aztec’s political organization, which has been described as hegemonic or indirect. Kingdoms conquered by the Aztecs remained independent in all internal affairs, their rulers were typically not removed, and representatives of the Aztec Emperor, such as provincial governors, were largely absent. I follow Murdock’s choice and use a value of 2 for the Aztec ethnic group through most of the paper; but I also subject the results to a robustness check where the Aztec group is assigned the maximum value of 4. Most other Amerindian groups are assigned a value of 0 or 1 in Murdock’s scale⁵, with the exception of a few groups organized in confederacies of city-states such as the Muisca of central Colombia or the Zapotecs of southern Mexico.

I combine the above variable with data on the ethnic structure of the population for each sub-national state in Latin America to construct a population-weighted average of Murdock’s Jurisdictional Hierarchy index for all states. The variable, which takes non-integer values between 0 and 4, is constructed using only the population and institutional data for Amerindian ethnic groups, and as such reflects the average level of institutional complexity among the Amerindian population residing in each state⁶.

It is worth pointing out that the construction of this measure of pre-colonial institutions requires the matching of ethnic groups listed under two different datasets. Indeed, the data on population shares comes from national censuses - which I accessed individually through each nation’s statistical agency. While a majority of ethnic groups receive the same name in the national censuses and in the Ethnographic Atlas, there are a number of cases where the names assigned in these two sources differ. I used a diversity of additional material

empire and who were not assigned a value of Jurisdictional Hierarchy in Murdock (1967). My results are not dependent on this choice.

⁵Please note that none of the ethnic groups in Latin America, as recorded by the Atlas, take a value of 3 of Murdock’s Jurisdictional Hierarchy index. Though, as emphasized previously, level and 3 and 4 are reserved to ethnic groups that were part of large states with several levels of intermediate bureaucracy between its ruler and the local community. Yet even that there are not ethnic groups with this level of institutional complexity, the construction of our index is still feasible since, as I will explain below, by using the population-weighted average of such index, its resulting values will therefore be capturing the average level of institutional complexity among all the Amerindian population living in each state, which will be taking non-integer values between 0 and 4.

⁶More explicitly, I construct a population-weighted average of Murdock’s Jurisdictional Hierarchy index for all ethnic groups living in each state by using the following formula:

$$PCI_{s,c} = \sum_e \frac{Ame_{e,s,c}}{Ame_{s,c}} \cdot \mathbf{JH}_e,$$

where PCI is our the state-level measure of my pre-colonial institution index in state s of country c . $Ame_{e,s,c}$ represents the total number of Amerindians of ethnic group e living in state s of country c . $Ame_{s,c}$ denotes the total number of Amerindians living in states s of country c . Finally, \mathbf{JH}_e denotes the respective level of institutional complexity of ethnic group e based on Murdock’s Jurisdictional Hierarchy index.

in order to make sure that as many ethnic groups as possible were matched - please refer to Table A.1 in the Appendix A for details. By this procedure I am able to match 102 Amerindian ethnic groups from the census data to the Atlas.

It is important to note as well that the data uses the most recent census available for each country, as previous versions would have a less comprehensive coverage of the Amerindian population (for instance not recording the exact ethnic group). It is worth mentioning that the percentage of Amerindians in the total population may change significantly between two censuses, as census questions are modified and social attitudes towards Amerindians evolve. While I cannot say much about how the results would be affected if different census years were used, I did try using two different censuses for Bolivia (2001 and 2012) as the data was sufficiently detailed in both cases. My results were similar in both cases.

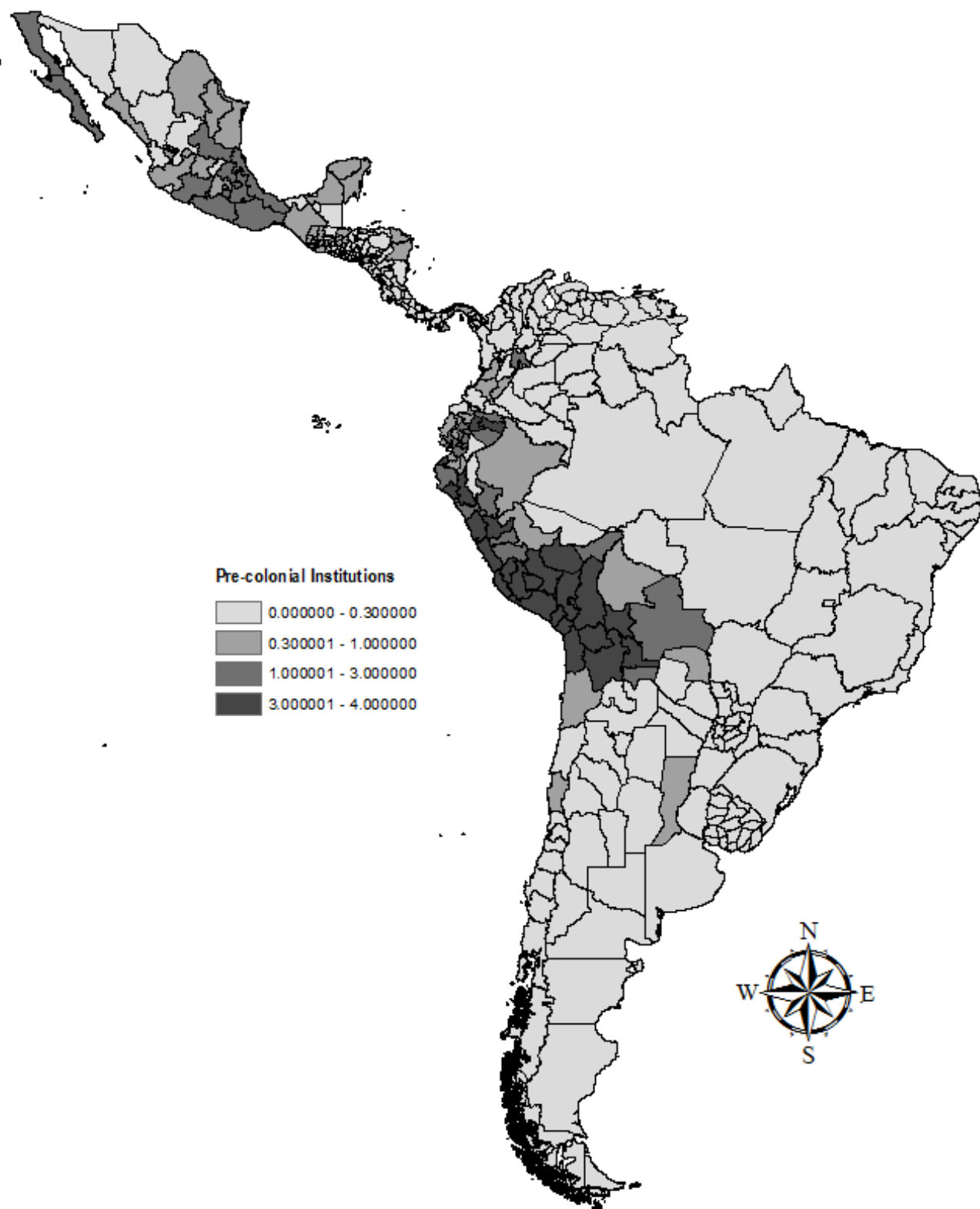
These 102 Amerindian groups for which a value of the index of Jurisdictional Hierarchy could be assigned represent 71% of the total Amerindian population of Latin America - albeit this percentage varies significantly from country to country. The fact that almost 30% of the Amerindian population could not be matched is to be expected given that the Ethnographic Atlas does not offer an exhaustive list of all groups but rather a survey of the groups for which anthropological work is available. For the Amerindian groups that could not be matched, I assign the minimum value of Jurisdictional Hierarchy under the assumption that small and less organized groups were more likely to remain unresearched by anthropologists. The assumption is supported by the fact that all groups present in the Atlas with a value of Jurisdictional Hierarchy equal to 1 or higher were matched to the census data. As a robustness check, I also experiment assigning non-matched groups a value equal to the average of all matched groups within the same state.

Figure 2.1 presents a visual overview of my measure of pre-colonial institutions across all subnational states in continental Latin America. The contours of the Inca and Aztec Empire are in evidence, as are the relatively advanced areas of the Yucatan peninsula and Central Colombia.

With the measure of pre-colonial institutions at hand, I investigate its influence on present-day socioeconomic outcomes in Latin America using the following econometric specification:

$$Y_{s,c} = \alpha_c + \beta PCI_{s,c} + \theta Amepop_{s,c} + \lambda \mathbf{X}_{s,c} + \varepsilon_{e,c} \quad (1)$$

Figure 2.1: Pre-Colonial Institutions in Latin America



In equation (1), $Y_{s,c}$ is an outcome variable such as a measure of schooling, health or economic well-being. Subscript c denotes subnational states, subscript s denotes countries, and α_c is a set of country-specific fixed effects. $PCI_{s,c}$ is the measure of pre-colonial institutions described above and $\mathbf{X}_{s,c}$ is a set of variables controlling for state characteristics such as population density, geography, and a number of colonial and pre-colonial factors potentially affecting socioeconomic outcomes. Finally, $Amerpop_{s,c}$ is the share of Amerindians in the total population of the state, a control variable that I single out for its importance.

Indeed, the share of Amerindians in the total population matters as Amerindians have traditionally suffered from ethnic discrimination, may experience difficulty integrating into a dominant non-Amerindian society (for example, because formal education may not be provided in their native language), and tend to be given a low priority by the national government. Under such circumstances, it should not be surprising that regions with a larger Amerindian population tend to be characterized by lower levels of socioeconomic outcomes (see Psacharopoulos and Patrinos (1994) for a detailed analysis of this issue).

Controlling for the share of Amerindians in the total population is important as institutional development typically went hand in hand with population density in the pre-industrial world. Thus, areas with more advanced pre-colonial institutions would also be areas with a larger Amerindian population, and therefore a larger share of Amerindians in the total population today. The positive effect of pre-colonial institutions that I hypothesize would be biased downwards if I do not control for the fact that more advanced areas will also tend to suffer more from discrimination towards Amerindians.

It is also worth pointing out that, given the nature of Murdock’s dataset, the variable $PCI_{s,c}$ may suffer from measurement error. This, however, would only result in a bias towards zero in the estimates of coefficient β . As most of the results rely on β being different from zero, I may say that conclusions would be stronger if this variable could be measured with more precision.

All our regressions include country fixed effects as these control for a wealth of characteristics shared by all states within the same nation. In particular, a number of colonial and post-colonial institutional factors will be common to all subnational units, such as the written laws and the constitution, the organization of public health and education, the balance of power between the different branches of the government, and so on. As I search to isolate the effect of pre-colonial institutions, accounting for as many colonial and post-colonial institutional factors as possible is important. Of course, it may still be the case that some of these factors display variability at the subnational level: some aspects of the law, for instance, may be applied more stringently in the capital city as compared

to far-off provinces. I cannot control directly for such effects, but note that if the way in which national institutions such as legal codes apply locally correlates with the local level of pre-colonial institutions, a potential explanation may be that more advanced pre-colonial institutions facilitate the operation of national law. In that case, the regressions would simply be capturing an additional indirect effect of pre-colonial institutions, one that works via a more effective functioning of the post-colonial state.

I will consider as dependent variables three indicators of education (percent of the population who completed primary education, percent who completed secondary education, average years of schooling), one indicator of public health (infant mortality rate), two indicators of economic well-being (percent of the population with access to drinking water, percent with access to electricity), and two indicators of overall economic development (GDP per capita and poverty rates)⁷. The battery of control variables at the state level will be discussed in the following section, as they are progressively introduced. The sources and precise definitions of all variables used in the paper can be found in Table A.2 in the Appendix A.

Before turning to our formal statistical analysis, Tables 2.1 to 2.2 offer an overview of the data. Table 2.1 lists all countries in the dataset together with their total population, the percentage of Amerindians in their population, and the percentage of their Amerindian population being matched to the Ethnographic Atlas. The data covers all countries in continental Latin American - a total of 17 countries and 324 subnational states.

It is worth pointing out that for Brazil the population shares of different Amerindian groups is only available at the level of regions (groups of 3 to 9 states). I assign to each Brazilian state the population shares of the region it belongs to. For Argentina the data is available at the state level but gives only a partial breakdown, with the population of only the main Amerindian groups of each state given. I complete the missing data for Argentina using national totals for each group and assumptions about the distribution of each group outside the states where they are most numerous. For Uruguay I do not have data on different Amerindian groups, only the population share of all Amerindians in each state. This, however, is not a problem for the construction of my measure of Pre-Colonial Institutions for Uruguay as we know that all Amerindian groups in Uruguay have a value of zero of Jurisdictional Hierarchy. For all other 14 countries I have a complete dataset giving population shares for all Amerindian groups in every state.

⁷Please note that, unless otherwise specified, most of the data for my dependent variables uses the most recent census available for each country. The time of each census varies, yet it ranges between 2002 and 2012. To find out the specific year of each census, please go to <http://redataam.org/redbin/RpWebEngine.exe/Portal?lang=eng>

Table 2.1: Latin American countries and their Amerindian populations

Country	Total population	Amerindian population as % of total population	Amerindian matched to Atlas as Amerindian population	population Ethnographic % of total	Number of states
Argentina	40,117,096	2.8%	25%		24
Bolivia	10,059,856	41%	84%		9
Brazil	190,755,799	0.4%	38%		27
Chile	15,116,385	4.6%	96%		13
Colombia	41,174,853	3.4%	47%		33
Costa Rica	4,301,712	1.6%	26%		7
Ecuador	14,451,115	7%	42%		24
El Salvador	57,44,113	0.2%	15%		14
Guatemala	11,237,196	39%	62%		8
Honduras	6,076,885	6.3%	96%		18
Mexico	103,263,388	5.7%	77%		32
Nicaragua	5,483,447	8%	27%		17
Panama	3,405,813	12%	29%		12
Paraguay	5,163,198	1.7%	55%		18
Peru	27,412,157	15%	96%		25
Uruguay	3,285,877	2%	100%		19
Venezuela	27,225,775	2.8%	79%		24
TOTAL	514,274,665	5%	71%		324

Table 2.2: Descriptive Statistics

	Observations.	Mean	Std. Dev.	Min	Max
<i>Dependent variables:</i>					
Infant mortality rate	324	20.6	10.8	1.4	56.4
Years of schooling	319	5.98	2.13	1.03	11.45
Primary school achievement	324	0.81	0.16	0.29	0.96
Secondary school achievement	324	0.41	0.15	0.05	0.75
Drinking water	324	0.86	0.17	0.05	0.99
Electricity	324	0.84	0.19	0.03	0.99
PPP GDP per capita	300	5763.52	4722.30	1248.88	40449.09
Poverty rate	272	29.18	20.78	1.23	81.67
<i>Main variable of interest:</i>					
Measure of pre-colonial institutions	324	0.57	1.05	0	3.99
<i>Baseline control variables:</i>					
Share of Amerindians in total population	324	0.11	0.19	0.00001	0.96
Population density	324	394.9	3407.41	0.13	58706.88
Land Suitability Index	321	0.56	0.29	0.002	0.998
Malaria Stability Index	321	1.3	1.4	0	5
Latitude	324	16.02	10.73	0.015	54.33
Altitude (<i>km.</i>)	324	0.68	0.92	0	4.33
Temperature (<i>Celsius</i>)	319	20.72	5.28	4.7	27.77
Land area (<i>sq. km.</i>)	324	63815.63	151115.3	44	1559162
Landlocked dummy	324	0.54	0.49	0	1
Distance to capital (<i>km.</i>)	324	464.08	477.69	0	2559.34
Inverse distance to coast	320	0.89	0.1	0.54	0.99
Oil & Gas dummy	324	0.16	0.36	0	1
Gold & Silver dummy	324	0.12	0.32	0	1
Other mines dummy	324	0.23	0.42	0	1

Table 2.3: Matrix of correlations

	Pre-Colonial Institutions	Share of Amerindian pop.	Infant Mortality Rates	Years of Education	Primary Education	Secondary Education	Drinking Water	Electricity	GDP per capita	Poverty Rates
Pre-Colonial Institutions	1									
Share of Amerindian pop	0.35	1								
Infant Mortality Rates	0.02	0.36	1							
Years of Education	0.05	-0.34	-0.75	1						
Primary Education	0.15	-0.33	-0.47	0.66	1					
Secondary Education	0.27	-0.17	-0.57	0.65	0.64	1				
Drinking Water	-0.19	-0.53	-0.41	0.41	0.33	0.24	1			
Electricity	-0.16	-0.55	-0.59	0.60	0.59	0.59	0.64	1		
GDP per capita	-0.10	-0.33	-0.54	0.51	0.44	0.56	0.40	0.59	1	
Poverty Rates	0.19	0.39	0.21	-0.05	-0.23	-0.12	-0.44	-0.39	-0.52	1

Table 2.2 presents descriptive statistics for the main variables used in the analysis, while Table 2.3 reports all bivariate correlations between our measure of pre-colonial institutions, the share of Amerindians in the total population, and all the outcome variables. I note that the correlation between pre-colonial institutions and the share of Amerindians in the population is positive but not that high (0.35) - proof that these two variables are not measuring the same phenomena. I may also note that while the share of Amerindians is negatively related to all our measures of socioeconomic success (negative correlations except for infant mortality and poverty rates, where higher values denote worse outcomes), this is by far not the case for pre-colonial institutions. The bivariate relationship between pre-colonial institutions and socioeconomic outcomes can be mildly positive, as reported in Table 2.3.

2.4 Empirical analysis

2.4.1 Baseline results

I begin the statistical analysis with a set of baseline regressions reported in Table 2.4. All regressions include country fixed effects and control variables are added progressively in order to appreciate their effect on the coefficient of interest. The dependent variable I select for this initial analysis is the percentage of the population who finished secondary schooling, in logarithmic form⁸. Standard errors are clustered at the country level throughout the chapter.

The first column of Table 2.4 reports the most simple regression where only the measure of pre-colonial institutions is included alongside fixed effects. The coefficient on pre-colonial institutions is already statistically significant at the 5% level and takes a value of 0.0378. The coefficient's magnitude doubles to 0.0894 and its statistical significance increases to 1% in the second column, where the share of Amerindians in the total population of each state is added. As expected, this last variable has a strong negative influence on the outcome measure and its absence was responsible for a negative bias in the effect of pre-colonial institutions. We are thus confirming previous research as to the overall poorer socioeconomic outcomes of the Amerindian population, but adding a new result whereby areas where Amerindians groups were characterized by higher pre-colonial institutional quality have better outcomes.

⁸Please note that the reason to use secondary education as dependent variable in table 2.4 is mainly to illustrate how the effect of pre-colonial institutions changes (or remains unchanged) as additional controls are added. I have run the set of regressions in table 2.4 for each of the eight dependent variables at my disposal and results are similar. In short, education is regarded as one among several indicators of socioeconomic development -and not our main channel of interest.

Table 2.4: Baseline results

Dependent variable: Percent of population having completed Secondary education (in logs)						
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-Colonial Institutions	0.0378** [0.0148]	0.0894*** [0.0214]	0.0751*** [0.0237]	0.0846*** [0.0238]	0.0877*** [0.0223]	0.0841*** [0.0220]
Share of Amerindian population		-0.746*** [0.192]	-0.632** [0.222]	-0.583** [0.234]	-0.594** [0.230]	-0.578** [0.231]
Log population density			0.0279** [0.0105]	0.0299** [0.0135]	0.0493*** [0.0151]	0.0541*** [0.0118]
Latitude				0.00954** [0.00344]	0.0107*** [0.00315]	0.0105*** [0.00305]
Malaria Stability Index				-0.0117 [0.0155]	-0.0159 [0.0139]	-0.0146 [0.0132]
Temperature (Celsius)				0.00284 [0.00559]	-2.41e-05 [0.00524]	0.00110 [0.00505]
Altitude (km.)				-0.0359** [0.0154]	-0.0341* [0.0171]	-0.0359* [0.0189]
Land area (sq. km.)					5.45e-08 [9.63e-08]	3.49e-08 [9.35e-08]
Landlocked dummy					-0.0652 [0.0552]	-0.0582 [0.0528]
Distance to capital (km.)					7.83e-05** [2.79e-05]	6.67e-05 [3.87e-05]
Inverse distance to coast					-0.895*** [0.306]	-0.855** [0.308]
Land Suitability Index						-0.0695 [0.126]
Oil & Gas dummy						0.0132 [0.0293]
Gold & Silver dummy						0.0530 [0.0436]
Other mines dummy						0.00345 [0.0320]
Country Fixed Effects	YES	YES	YES	YES	YES	YES
Observations	324	324	324	317	317	317
Adjusted R-squared	0.703	0.767	0.775	0.781	0.789	0.788

Notes: Cluster standard errors at country level are in brackets. *** p<0.01, ** p<0.05, * p<0.1

The next four columns of Table 2.4 add a large number of state-specific characteristics which may have an effect on socioeconomic outcomes and whose absence could create an omitted variable bias. In column 3, I control for the present-day population density of each state - as the provision of education may be more costly in less densely settled territories. As expected, areas of higher population density tend to have better outcomes, but the effect of pre-colonial institutions continues to be large and statistically significant. It is important to note that present-day population density may be potentially be considered as a “bad control” as pointed out by my priors Michalopoulos and Papaioannou (2013). For the case of Latin America, while it is true that European colonizers often settled in the areas of high pre-colonial development, subsequent historical events (especially post-independence) meant that regions where no advanced pre-colonial group existed became highly populated. A few examples are the regions around the cities of Sao Paulo, Rio de Janeiro, Buenos Aires or Santiago: among the richest and most densely populated in Latin America yet our index of pre-colonial development is essentially zero for all of them. Yet I have experimented dropping this variable from our control set and the coefficients of interest are essentially unchanged. I am therefore quite confident that bad controls are not biasing our results.

Columns 4 to 6 deal with the important issue of geography, including aspects such as climate, location and natural resources. It is natural to think that more advanced pre-colonial groups would have displaced rivals and established themselves in regions characterized by more fertile land, better access to rivers, a milder climate, etc. If that was the case, their present-day advantage may be due to those geographic factors, and not to their superior institutions. In addition to this, and working in the same direction, less advanced groups would have been less able to resist losing their land and other resources when the colonial conquest reached their territory.

With this in mind, column 4 augments the baseline regression with four indicators of climate: latitude, altitude, temperature, and an index of Malaria prevalence (higher values denote higher incidence). Column 5 includes the area of the state in question plus three indicators of its locational advantage: distance to the capital, distance to the sea, and a dummy for landlocked states. Column 6, finally, directly measures the most important forms of natural resource wealth by adding an index of land suitability for agriculture and dummy variables indicating the presence of oil or gas fields, gold or silver mines, and any other mines.

The main result of these three columns is that the coefficient on pre-colonial institutions remains statistically significant at the 1% level in all cases and its magnitude is not much affected. In column 6, when all controls are included, the coefficient takes a value of 0.0841. As the dependent variable is measured in logarithmic form, this coefficient indicates that

an increase in the average level of pre-colonial institutions by 1 unit is associated with an increase in secondary school achievement of around 8%. This is a large effect when we consider an average value of secondary school achievement across all states of 41% with a standard deviation of 15% - passing from a pre-colonial population of tribesmen to one of multi-city chiefdoms (increase of 2 units) would lead to a one standard deviation change in secondary schooling.

Turning to our state-specific indicators of geographic advantage, latitude, altitude, distance to the capital and distance to the sea all appear to have a statistically significant relationship with the outcome variable. Latitude and altitude seem to pick up the effect of all climatic factors, as neither malaria prevalence nor temperature are statistically significant in their presence. Somewhat surprisingly, distance to the capital is positively related to education while being landlocked appears to carry no additional penalty once distance to the coast has been accounted for. Finally, none of the four indicators of natural resource wealth exerts a statistically significant effect on secondary school achievement. Arguably, the gains from having these resources at hand are counteracted by opposite effects much discussed in the ‘natural resource curse’ literature.

The results of Table 2.4 may be reproduced over the full array of socioeconomic indicators, as shown in Table 2.5. This Table takes as its baseline the regression reported in the last column of Table 2.4, with all state characteristics considered so far, and considers as dependent variable each of the eight outcome measures at my disposal. Remarkably, the index of pre-colonial institutions is consistently related with better outcomes for all of them: it is positively related with measures of education, drinking water, electricity and GDP per capita, and negatively related with infant mortality and poverty rates. In all cases the relationship is statistically significant at the 5% level or better.

The effect of pre-colonial institutions is not only statistically significant, the magnitude of the effect is also large. Since all dependent variables are used in logarithmic form, coefficients may be interpreted directly as semi-elasticities. Interestingly, the largest effects are observed for our measures of overall economic development. A 1-unit increase in the index of pre-colonial institutions is associated with a 20% increase in GDP per capita and a 12% decrease in the poverty rate. The effects for all other dependent variables are in the 3 to 8% range for a 1-unit increase, in all cases a sizeable change.

Table 2.5: Baseline results with 8 different dependent variables

Dependent variable	Infant Mortality (1)	Years of Schooling (2)	Primary education (3)	Secondary education (4)	Drinking water (5)	Electricity (6)	Log GDP per capita (7)	Poverty rates (8)
Pre-Colonial Institutions	-0.0594*** [0.0184]	0.0568** [0.0201]	0.0276** [0.0115]	0.0841*** [0.0220]	0.0535*** [0.0170]	0.0858*** [0.0290]	0.199** [0.0848]	-0.117*** [0.0371]
Controls included:								
Share of Amerindian pop.	YES	YES	YES	YES	YES	YES	YES	YES
Log population density	YES	YES	YES	YES	YES	YES	YES	YES
Controls for geography, location and natural resources	YES	YES	YES	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	317	316	317	317	317	317	307	270
Adjusted R-squared	0.787	0.816	0.875	0.788	0.503	0.651	0.623	0.848

Notes: Cluster standard errors at country level are in brackets. *** p<0.01, ** p<0.05, * p<0.1

2.4.2 Controlling for colonial activities

While the results so far control for institutional factors which are constant at the national level and for a vast array of geographic characteristics, other historical factors not yet considered could be a source of bias. In particular, pre-colonial institutional development may be related to the colonial activities put in place following the conquest. Indeed, the territories of the most advanced pre-colonial civilizations - central and southern Mexico, the Andes - were also the source of most Amerindian labour. The availability of this labour made possible a range of economic activities during the colony, most notably silver mining and the agricultural latifundia. If these activities then have an effect on present-day outcomes, pre-colonial institutions would be correlated with socioeconomic development but the causal mechanism would work through the colonizing process.

To test for this alternative explanation I take advantage of the work of Bruhn and Gallego (2012), who investigate the role of colonial activities on economic development in Latin America using states as the unit of analysis. They classify states into four mutually exclusive groups according to the main economic activity taking place in their territory during the colonial period. These four groups are:

- a) Mining. In particular the gold mines of Brazil, the silver mines of Mexico, Peru and Bolivia, and the associated mines producing mercury for the process of silver extraction through amalgamation.
- b) Plantations. Places dedicated to the cultivation of high-value cash crops for the export market, in particular sugar, tobacco and cotton. Plantations relied essentially on slave labour.
- c) Other colonial activities. Places where the dominant economic activity was agricultural production for the local market (from Amerindian lands or from latifundia) and industry.
- d) No colonial activities. Places where the colonial state had marginal or no influence, like remote parts of the Amazonian rainforest and the extreme south of Argentina and Chile.

It is important to note that Bruhn and Gallego (2012) don't use these four groups in their analysis. Instead, they combine the information on the type of economic activity in each state with data on pre-colonial population density to produce a classification into three types of colonial activities which they refer to as "bad", "good" and "ugly". I don't follow their approach as it incorporates value judgements as to what is believed to be "good" or "bad" (let alone "ugly"). The classification of colonial activities into mining, plantations, and others is much less likely to be affected by our own beliefs.

I incorporate dummy variables for the first three types of economic activities, leaving the case of no colonial activities as our excluded category. Results are reported in Table 2.6, where all regressions include country fixed effects and our full range of baseline control variables⁹.

Table 2.6 is strongly supportive of our thesis. Indeed, the measure of pre-colonial institutions continues to have a positive and statistically significant effect on the eight dependent variables I consider. The size of the coefficients is not much affected with respect to Table 2.5, only the effects on educational achievement and electricity provision are somewhat reduced. This indicates that the relationship between pre-colonial institutional development and present-day outcomes is largely not mediated by the type of economic activity put in place during the colony.

Turning to the effects of colonial activities on present-day outcomes, Table 2.6 gives us a mixed picture. The effect seems clearest on overall measures of economic development, as states associated with mining and plantation agriculture have lower levels of GDP per capita than states left untouched by the colonial economy. This is in line with Bruhn and Gallego (2012), who based most of their analysis on the effects on GDP per capita. For other measures of socioeconomic development, however, the evidence is less conclusive. Areas where slave-based plantations were located are indeed characterized by lower secondary education and less access to drinking water, albeit the relationship only marginally achieves statistical significance. But no further statistically significant effects are estimated for areas formerly dedicated to mining or other colonial activities. Overall, while colonial activities may well play a role in determining current development outcomes, the results show that their consideration does not diminish the importance of pre-colonial institutions.

⁹Please note that this set of controls on colonial activities may also be regarded as "bad controls". I am sensitive to this potential problem but do not believe it concerns most of the variables on colonial activities. A "bad control" is a variable which is itself an outcome of the regressor of interest (in this case pre-colonial institutions). This is poorly applied to the two most important types of colonial activities, plantation agriculture and mining. Plantation agriculture was largely determined by climate conditions, like the suitability to grow sugarcane, while mining was largely determined by the existence of mine fields.

On the other hand, I would agree that the dummy for "other colonial activities", which includes industry, may plausibly be regarded as a bad control. I have experimented dropping this variable from our control set and the coefficients of interest are essentially unchanged. As with my control on current population density, I am quite confident that bad controls are not biasing my results.

Table 2.6: Controlling for colonial activities

Dependent variable	Infant Mortality (1)	Years of Schooling (2)	Primary education (3)	Secondary education (4)	Drinking water (5)	Electricity (6)	Log GDP per capita (7)	Poverty rates (8)
Pre-Colonial Institutions	-0.0601*** [0.0194]	0.0430** [0.0157]	0.0176** [0.00653]	0.0675*** [0.0181]	0.0457** [0.0207]	0.0559** [0.0226]	0.197** [0.0819]	-0.118** [0.0428]
Other colonial activities	0.0405 [0.0492]	-0.0227 [0.0269]	-0.00777 [0.00868]	-0.0357 [0.0360]	-0.0480 [0.0322]	-0.0243 [0.0190]	-0.150 [0.113]	0.140 [0.121]
Mining colonial activities	0.00649 [0.0697]	0.00453 [0.0488]	0.0164 [0.0103]	0.0155 [0.0598]	-0.0461 [0.0443]	-0.00885 [0.0218]	-0.309** [0.119]	0.279** [0.115]
Plantation colonial activities	0.157 [0.0898]	-0.0827 [0.0602]	-0.0153 [0.0205]	-0.127* [0.0646]	-0.0466* [0.0262]	-0.0625 [0.0453]	-0.341** [0.154]	0.335 [0.225]
Controls included:								
Share of Amerindian pop.	YES	YES	YES	YES	YES	YES	YES	YES
Log population density	YES	YES	YES	YES	YES	YES	YES	YES
Controls for geography, location and natural resources	YES	YES	YES	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	280	279	280	280	280	280	280	270
Adjusted R-squared	0.772	0.819	0.914	0.815	0.472	0.735	0.598	0.852

Notes: Cluster standard errors at country level are in brackets. *** p<0.01, ** p<0.05, * p<0.1

2.4.3 Controlling for other pre-colonial characteristics

If the results so far clearly point towards a persistent role of pre-colonial institutions on current socioeconomic development, one may still argue that pre-colonial features other than institutional complexity may explain the results. As I mentioned briefly, institutional complexity usually correlates with economic development, and it is possible that richer pre-colonial societies were able to adapt better and take advantage of the new colonial environment simply because of their wealth. Furthermore, the Ethnographic Atlas provides a large array of cultural and economic practices of the societies it surveys. I am therefore in a position to control for a number of pre-colonial characteristics other than the complexity of their political structure - and I do so in what follows.

I start with overall economic development in pre-colonial times. Clearly, measures of income per head are not available for this time period in the Americas, but I may follow much of the relevant literature and rely on estimates of population density as a proxy for overall economic development (see, for instance, Acemoglu et al. 2002). The data on pre-colonial population density at the state level comes from Bruhn and Gallego (2012), and Table 2.7 adds this variable as an additional control to the regressions reported in Table 2.6.

Once again, the results are fully consistent with the thesis of the present Chapter. The coefficient of pre-colonial institutions is hardly affected by the inclusion of this variable and remains statistically significant for all dependent variables. Perhaps surprisingly, pre-colonial population density has a statistically significant effect on only one of the eight dependent variables I consider - GDP per capita. This suggests that whatever effect pre-colonial economic development may have on present-day socioeconomic outcomes, the transmission mechanism works mainly through institutional persistence. I also note that I find a negative effect of pre-colonial population density on GDP per capita, as in the "reversals of fortune" thesis by Acemoglu et al. (2002). For alternative views on the reversals of fortune thesis in the Americas see Chanda et al. (2014) and Maloney and Valencia (2015).

Table 2.7: Controlling for pre-colonial characteristics: population density

Dependent variable	Infant Mortality (1)	Years of Schooling (2)	Primary education (3)	Secondary education (4)	Drinking water (5)	Electricity (6)	Log GDP per capita (7)	Poverty rates (8)
Pre-Colonial Institutions	-0.0593*** [0.0184]	0.0438** [0.0158]	0.0182** [0.00647]	0.0696*** [0.0186]	0.0442** [0.0192]	0.0553** [0.0226]	0.207** [0.0780]	-0.123** [0.0459]
Pre-colonial population density	-0.00604 [0.0220]	-0.00592 [0.00757]	-0.00437 [0.00432]	-0.0157 [0.0112]	0.0114 [0.0115]	0.00391 [0.00803]	-0.0704* [0.0352]	0.0352 [0.0249]
Other colonial activities	0.0443 [0.0558]	-0.0190 [0.0263]	-0.00501 [0.00902]	-0.0258 [0.0317]	-0.0552 [0.0385]	-0.0268 [0.0197]	-0.106 [0.105]	0.118 [0.115]
Mining colonial activities	0.00914 [0.0711]	0.00710 [0.0484]	0.0183 [0.0106]	0.0224 [0.0565]	-0.0511 [0.0492]	-0.0106 [0.0213]	-0.279** [0.114]	0.264** [0.112]
Plantation colonial activities	0.164 [0.0968]	-0.0756 [0.0617]	-0.0100 [0.0217]	-0.108 [0.0638]	-0.0602 [0.0351]	-0.0671 [0.0483]	-0.257* [0.144]	0.293 [0.223]
Controls included:								
Share of Amerindian pop.	YES	YES	YES	YES	YES	YES	YES	YES
Log population density	YES	YES	YES	YES	YES	YES	YES	YES
Controls for geography, location and natural resources	YES	YES	YES	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	280	279	280	280	280	280	280	270
Adjusted R-squared	0.771	0.819	0.914	0.816	0.472	0.734	0.613	0.853

Notes: Cluster standard errors at country level are in brackets. *** p<0.01, ** p<0.05, * p<0.1

In Table 2.8, I take an additional step and control for nine social and economic characteristics of pre-colonial societies other than their institutional complexity. These characteristics are the fraction of the population dedicated to gathering, hunting, fishing and agriculture; their typical pattern of settlement (from fully nomadic to compact and permanent settlements); their degree of class stratification; a dummy for the existence of slavery; a dummy for the existence of elections in determining leader succession and, finally, a dummy for the existence of inheritance rules for property (see Table A.3 in the Appendix A for detailed definitions). For each of them, I proceed as for the measure of pre-colonial institutions: I calculate the population-weighted average among all Amerindian groups present in the state¹⁰. The first four measures, all relating to the economic activity of the population, are included simultaneously in column 2.¹¹ All other variables are included separately in the remaining columns of the Table. The regressions also control for the different colonial activities as in Table 2.6 and for pre-colonial population density as in Table 2.7, besides all the state-specific characteristics and country fixed effects that have been included all along. The dependent variable is the percentage of the population with secondary education.

As Table 2.8 makes clear, the inclusion of these additional pre-colonial characteristics does not challenge the importance of pre-colonial institutions. In all regressions the coefficient on the measure of pre-colonial institutions remains positive and statistically significant at the 1% level. The magnitude of the coefficient is remarkably consistent, fluctuating closely around the value of 0.0700 in all but one case (column 2, where the coefficient equals 0.0963). Thus, the coefficient is usually very similar to what is obtained before any of these additional pre-colonial characteristics is controlled for (first column of Table 2.8). In most cases, the additional pre-colonial characteristics considered turn out to have no statistically significant effect on education. This reinforces the thesis, which regards institutional complexity as the crucial aspect of pre-colonial societies affecting current outcomes.

The exercise of Table 2.8 may be reproduced using the other seven dependent variables considered previously. While I do not report these results for conciseness, I have carried them out and the importance of pre-colonial institutions is never challenged. The sign and statistical significance of pre-colonial institutions carries through for all seven alternative outcome variables and in essentially all specifications considered in Table 2.8.

¹⁰For the Amerindian groups that could not be matched to the Atlas I assign a value equal to the average value of all other groups within the state. This is different from what I did for my measure of pre-colonial institutions, as the variables considered here are not necessarily related to social complexity, and could not be assumed to take the lowest possible value. Uruguay is excluded from table 2.8 as I don't have enough data to calculate these additional variables for it.

¹¹These four variables do not sum up to 1, as a fraction of the population may be counted in more than one of them, and sometimes in none of them.

Table 2.8: Controlling for pre-colonial characteristics: socioeconomic factors

Dependent variable: Percent of population having completed Secondary education (in logs)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pre-Colonial Institutions	0.0705*** [0.0188]	0.0963*** [0.0193]	0.0712*** [0.0190]	0.0719** [0.0249]	0.0705*** [0.0188]	0.0718*** [0.0201]	0.0715*** [0.0208]
Pre-colonial population density	-0.0231 [0.0139]	-0.0210 [0.0120]	-0.0237 [0.0139]	-0.0231 [0.0138]	-0.0231 [0.0139]	-0.0244 [0.0138]	-0.0230 [0.0138]
Other colonial activities	-0.0360 [0.0359]	-0.0455 [0.0383]	-0.0387 [0.0372]	-0.0367 [0.0375]	-0.0357 [0.0383]	-0.0380 [0.0353]	-0.0367 [0.0368]
Mining colonial activities	0.0139 [0.0561]	0.00217 [0.0501]	0.0140 [0.0554]	0.0130 [0.0520]	0.0143 [0.0616]	0.0112 [0.0543]	0.0135 [0.0568]
Plantation colonial activities	-0.111 [0.0709]	-0.111 [0.0700]	-0.120 [0.0733]	-0.112 [0.0759]	-0.111 [0.0727]	-0.115 [0.0705]	-0.112 [0.0721]
Population employed in: Gathering		0.0304 [0.489]					
Hunting		0.424 [0.513]					
Fishing		0.625** [0.222]					
Agriculture dependence		-0.458*** [0.134]					
Settlement pattern			-0.0036 [0.00402]				
Class Stratification				-0.0025 [0.0237]			
Slavery					0.0024 [0.0597]		
Election						-0.115 [0.0827]	
Property rights							-0.0209 [0.0570]
Controls included:							
Share of ethnic population	YES	YES	YES	YES	YES	YES	YES
Log population density	YES	YES	YES	YES	YES	YES	YES
Controls for geography, location and natural resources	YES	YES	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES	YES	YES
Observations	261	261	261	261	261	261	261
Adjusted R-squared	0.810	0.821	0.810	0.809	0.809	0.811	0.810

Notes: Cluster standard errors at country level are in brackets. *** p<0.01, ** p<0.05, * p<0.1

2.4.4 Robustness checks

I have carried out a number of robustness checks on the above results, some of which I have already referred to:

- I assigned to all non-matched Amerindian groups from our census data a value of the index of Jurisdictional Hierarchy equal to the average value of all matched groups within the state (instead of a value of 0).
- I assign the maximum value of 4 as the Jurisdictional Hierarchy of the Aztec ethnic group.
- I exclude Brazil and Argentina from our regressions, as these two countries do not have complete data on the ethnic composition of their population at the state level.

In all cases, the results carry through all these checks unchallenged with only minor quantitative changes in the estimated coefficients.

2.4.5 Comparing rural and urban regions

As a final piece of evidence, I have gathered data allowing me to run separate regressions for the rural and urban regions of Latin American at the level of subnational states. While most of the data is not available at this level of disaggregation, I was able to find separate values for the rural and urban region of each state for four dependent variables (primary education, secondary education, access to drinking water, access to electricity) and for the ethnic composition of the population, which allows me to calculate the percentage of Amerindians in the total population and to construct my measure of pre-colonial institutions. All other control variables may be used in the analysis, but their values do not change between the rural and urban area of any given state. Argentina is omitted for this exercise, as there is no information on the distribution of its Amerindian population between urban and rural areas.

The interest of this exercise is that, in accordance with the discussion in section 2.2 of this Chapter, the importance of pre-colonial institutions ought to be far more marked in rural areas. Amerindians may be numerous in urban areas, but by migrating to them they enter a process of cultural assimilation within the dominant mestizo society. Amerindians no longer rely on their pre-colonial institutions once they leave their rural communities, as a different set of institutional arrangements is imposed upon them. If my hypothesis is correct, I should find that the positive relationship between pre-colonial institutions and socioeconomic development is stronger among rural areas.

Table 2.9: Contrasting rural and urban areas

Dependent variable	Primary education		Secondary education		Drinking water		Electricity	
	rural	urban	rural	urban	rural	urban	rural	urban
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pre-Colonial Institutions	0.0256** [0.0111]	0.0107** [0.00475]	0.0920*** [0.0256]	0.0340** [0.0131]	0.0846** [0.0381]	0.0370 [0.0256]	0.128*** [0.0401]	0.00567 [0.00578]
Controls included:								
Share of Amerindian pop.	YES	YES	YES	YES	YES	YES	YES	YES
Log population density	YES	YES	YES	YES	YES	YES	YES	YES
Controls for geography, location and natural resources	YES	YES	YES	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	291	290	291	290	291	290	291	290
Adjusted R-squared	0.881	0.933	0.815	0.886	0.615	0.226	0.693	0.617

Notes: Cluster standard errors at country level are in brackets. *** p<0.01, ** p<0.05, * p<0.1
Argentina has been removed from the sample for this table.

And indeed, the results clearly support this prior. Table 2.9 reproduces the regressions of Table 2.5, where the set of control variables includes country fixed effects, population density, and measures of climate, location and natural resources. I consider the four dependent variables mentioned above, and for each case run separate regressions using all rural areas or all urban areas. As it turns out, the effect of pre-colonial institutions is positive and statistically significant for the four cases covering rural areas, while only two of the four cases covering urban areas reach statistical significance. More important, the coefficient on pre-colonial institutions is always much larger for rural areas - between two and three times larger in most cases¹². As an example, it takes a value of 0.0920 for rural areas when the dependent variable is secondary education as opposed to 0.0340 for the corresponding regression using urban areas. I conclude that these results further solidify the hypothesis of the present Chapter.

2.5 Concluding remarks

If one thing has been learned from the last two decades of research on economic development over the very long run it is that the past cannot be easily cast aside. Every society builds on the successes and mistakes of its predecessors, and inherits a set of rules and institutions that are usually modified only gradually. While this seems obviously true

¹²To further test whether this difference is statistically significant between areas, I also experimented developing an empirical strategy where both urban and rural areas are included together into a single regression. I then interacted our index of pre-colonial institutions -as well as the set of controls- with a dummy on areas -where rural areas will be equal to 1 and urban zero. Since our reference category is urban areas, the effects of pre-colonial institutions in rural areas will be captured by the interaction between the index of pre-colonial institutions and the dummy on areas. In all but one of the coefficients of this interaction are positive and statistically significant. Results are reported in Table A.4 of Appendix A.

for the "*winners*" of economic history, the European nations that colonized the world, it is also the case for the "*losers*", those nations being colonized. What came out of the colonizing process throughout the world was not a mirror image of European society but a new reality where pre-colonial culture and institutions survived, often below a layer of official or dominant culture. These two layers interact and modify each other, and both of them ought to be considered in the study of today's developing countries.

The present Chapter brings support to the above assertions, and contributes to the substantive evidence already in place for the case of Africa. As the empirical results show, Latin American pre-colonial institutions - and more precisely the degree of political complexity - are powerful predictors of present-day measures of socioeconomic development. Several aspects render the evidence particularly convincing. First, the results are obtained controlling for country fixed effects, thus factoring out many institutional factors playing a role at the national level. Second, I introduce a large array of controls for geographic factors including climate, location and the presence of natural resources. Third, I consider additional historical forces such as the type of economic activity in place during the colony and the economic and social profile of pre-colonial societies (besides their institutional complexity). Finally, I show how the influence of pre-colonial institutions is far stronger in rural areas, which is in accordance with the historical account I give for the transmission of pre-colonial factors.

The present Chapter, together with the literature it contributes to, enhances our understanding of how developing countries got to where they are now. Understanding this is important in its own right, but also increases the chances of making the right decisions when considering where they head to in the future.

Chapter 3

Pre-Colonial Institutions, Ethnic Homelands, and Economic Development in Latin America

Abstract

This Chapter studies the long-run effects of pre-colonial institutions on present-day economic development in Latin America on the basis of the historical ethnic homelands. The main hypothesis is that ethnic homelands inhabited by more advanced ethnic groups –as measured by their levels of institutional complexity– relate to better economic development today. To track these long-run effects, I construct a new dataset by digitising historiographical maps allowing me to pinpoint the geospatial location of ethnic homelands as of the XVI century. As a result, 375 ethnic homelands are created. I then capture the levels of economic development at the ethnic homeland level by making use of alternative economic measures –satellite light density data. After controlling for country-specific characteristics and applying a large battery of geographical, locational, and historical factors, I found that the effects of pre-colonial institutions relate to a higher light density –as a proxy for economic activity– in ethnic homelands where more advanced ethnic groups lived.

3.1 Introduction

Very little is known within the economic history literature about the role of ethnic groups that lived before colonisation in the creation of wealth of the present-day Latin America's nations. Over the past decade, though, economists have placed enormous efforts in studying the long-run effects of ethnic-specific characteristics on today's economic development (Fenske 2014; Michalopoulos and Papaioannou 2013; Nunn and Wantchekon 2011).

These studies, however, have mainly examined the African case and most of their empirical examinations tend to be on the basis of the historical ethnic homelands. Interestingly, in searching for this link, the history of ethnic institutional advancements appears as one of the most powerful ethnic features that can explain over the long-run the differences of economic development.

In the previous Chapter of the present thesis, I explored empirically the association between the institutional complexity of ethnic groups and socioeconomic development outcomes in Latin America. The main results of this Chapter is that subnational states with a history of more advanced pre-colonial institutions relate to better levels of today's socioeconomic development outcomes -principally through their effects on the Amerindian population.

While the previous Chapter already uncovered the positive effects of pre-colonial institutions on present-day development in Latin America, the present Chapter addresses it in a rather different empirical fashion. As noted before, Chapter 2 uses the contemporary Latin American subnational states as the main unit of analysis, whereas in the present Chapter my main unit is now the ethnic homeland. Yet, both empirical investigation still allow me to include country fixed effects as a relevant method to reduce identification problems arising from historical shocks, cultural adaptations, or colonial and post-colonial institutions.

Thus, the main goal of this study is to empirically explore the link between pre-colonial institutions and contemporary development in Latin America through the effects of the historical ethnic homelands. To uncover this association, I construct a new dataset drawn from the series of Peter G. Murdock's historiographical publications of maps (1951, 1975) portraying the locations of ethnic homelands in Latin America as of the XVI century. I do so by digitising these maps allowing me to pinpoint the geospatial location of ethnic homelands in this region.

Using ethnic homelands as the main unit of analysis in the present empirical investigation may offer some methodological advantages and differences over the method applied in Chapter 2. First, the current study uses geospatial data on satellite light density at night as an alternative measure of economic development at this novel unit of analysis. It is worth pointing out that this unconventional measure of economic development is widely used in the existing literature addressing the long-run development; and the main idea for this is to overcome the paucity of the data at the ethnic homeland level (Michalopoulos and Papaioannou 2013).

Second, our index measuring the levels of pre-colonial institutions for the present investigation may be more accurate. As in Chapter 2, the current analysis uses Murdock's

(1967) index of *Jurisdictional Hierarchy beyond the Local Community* to measure the institutional complexity of ethnic groups -please refer to Chapter 2 for more details on this index. Though, unlike the method implemented in the previous Chapter, the index now does not need data on the population shares of the different Amerindian groups. As you may recall from the second Chapter, data on the share of the Amerindian groups was used in order to calculate for each subnational state the population-weighted average of Murdock's Jurisdictional Hierarchy index for all the Amerindian groups. Yet such population-weighted index did not cover all the Amerindian groups in each subnational state. On the other hand, the index for the current empirical investigation only uses the discrete value of Murdock's Jurisdictional Hierarchy index of the most representative ethnic group that lived in each ethnic homeland. Of course, the index may also suffer of measurement errors coming from the possible imprecisions in developing the historiographical maps. Nevertheless, it is worth experimenting both routes by which case one may find that the general hypothesis of the present doctoral thesis solidifies.

Third, I relax the assumptions of the main hypothesis. In Chapter 2, I assume that the pre-colonial institutions are mainly attached onto people -principally with Amerindian origins. Though, Amerindians may have changed considerably over the time as a result of diseases, destruction, exploitation and forced displacements during the colonial and post-colonial periods. Moreover, the fact that some subnational states today have a very low proportion of Amerindian population, one may also want to argue if such assumption is still robust. The present empirical investigation assumes, instead, that the pre-colonial institutions are attached onto geographical regions. The main idea is that as new people move into a given geographic region they are influenced by the institutional framework already in place there. This methodological strategy then allows me to only assume that regions where more advanced ethnic groups lived prior colonisation -as measured by their institutional complexity- tend to reflect as a whole better levels of economic development today -irrespective of the number of Amerindians groups that have resided or live there.

In order to attest accurately the long-run effects of pre-colonial institutions in Latin America, a substantial array of geographical and locational fundamentals are also added to the new dataset. For this, as with the main development outcome -satellite light density-, data drawn through geospatial methods is also generated at the ethnic homeland level. Data at this novel level of analysis allows me to account seriously the important role of key geographical and locational forces driving also the variation of our main economic development outcome.

Additionally, I supplement the new dataset with the state-level data that was prepared for the previous Chapter in an effort to include more historical information other than the levels of institutional complexity of ethnic groups. This allows me to control for

important historical factors (e.g. pre-colonial population density) that may be also behind the distribution of my main development outcome.

By developing a comprehensive methodological strategy, the present Chapter provides a deeper understanding of the formation of contemporary development over the long-run of today's nations in Latin America. The main results essentially support my previous findings. I show that ethnic homelands in Latin America that were historically inhabited by more advanced ethnic groups -as measured by their levels of institutional complexity- relate to a higher light density today -as a proxy for economic activity. These results hold even after controlling for various observable and hard-to-account-for country-specific features computed at the ethnic homeland level as well.

The rest of the Chapter is structured as follows. Section 3.2 provides the main links in the literature with the present research addressing the effects of ethnic homelands in Latin America on contemporary development. In section 3.3, I provide a very brief historical explanation given evidence as to why ethnic homelands may have affected development in Latin America over the long-run. Section 3.4 outlines the data and novel methodology implemented to uncover this association. Section 3.5 presents the main empirical results. Section 3.6 concludes.

3.2 Related literature

Within economics, the study of Michalopoulos and Papaioannou (2013) is the only formal contribution investigating empirically the relationship between ethnic institutional advancements and contemporary development through the effects of historical ethnic homelands. Yet the authors focus on Sub-Saharan Africa, where colonisation took a rather different form than in Latin America. Relying on the large variation of institutional complexities that ethnic groups had prior colonisation, they found a positive relationship between pre-colonial institutions and economic development outcomes -as measured by light density at night. My research advances Michalopoulos and Papaioannou's findings by analysing this link too but within the Latin American context, and on the basis of the historical ethnic homelands as well. The main findings of the present study are consistent to Michalopoulos and Papaioannou's hypothesis: ethnic homelands historically inhabited by ethnic groups with more complex institutional features are related to better levels of economic development in Latin America today.

This empirical investigation also provides evidence on the growing body of comparative development studies investigating the origins of wealth of the developing countries.

In the Latin American context, these studies offer several approaches but most of their conjectures rest on the role of colonial institutions shaping the paths of development in the long-run. (Hall and Jones 1999; Acemoglu et al. 2001; Banerjee and Iyer 2005; Sokoloff and Engerman 2011). My research expands these findings by looking at the variation of the institutional complexities of ethnic groups in the advent of the Europeans' arrival in the Americas in the XVI century.

The line of research of the present paper is also linked to an influential literature addressing the persistence of institutional factors (North 1990; Acemoglu et al. 2001; Banerjee and Iyer 2005; Angeles 2011; Dell 2010; Oto-Peralías and Romero-Ávila 2016); state capacity of nations (Diamond 1997; Acemoglu et al. 2013); and state formation (Bockstette et al. 2002). My results offer novel evidence on the important role of institutional advancements of ethnic groups in shaping the paths of economic development over the long-run in Latin America.

3.3 Historical context

The following section offers a brief historical context placing emphasis on the persistency of institutional advancements of ethnic groups affecting economic development over the long-run in Latin America.

Certainly, various reasons would make almost implausible in using traditional ethnic homelands in Latin America to explore the long-run effects of ethnic features on present-day development outcomes. No single ethnic group in Latin America escaped from the various havocs introduced through colonisation. Diseases, destruction and exploitation transformed the life and space of the large variety of ethnic groups that lived prior colonisation. The map of the Americas was then drawn based on the large delimitations marked by Europeans overlooking the existence of numerous ethnic homelands. However, as has been noted before, colonisation expanded under different paths depending on major geographical determinants, as well as on key ethnic-specific characteristics that the Europeans encountered at their arrival in the Americas in 1492.

In the advent of colonisation, detailed historiographical evidence suggests that several ethnic groups lived within their own ethnic homeland all along Latin America. These ethnic groups had a rather different way of organisation. As such, a few of them had established large centralised political systems in Central Mexico and some parts in South America. For instance, in South America, the Incas implemented the largest and more sophisticated political and administrative systems in the Americas (e.g. *corvée* labour or

mita). In Central Mexico, the Aztecs established institutions rivalling those in the Inca Empire. Some other groups located, for instance, in what is today Colombia and Ecuador had also similar paramount chiefdom characteristics (e.g. Chibchas, Quito or Catis) for which systems large enough to control various kingdoms beyond the fringe of their local communities were one of their singular ethnic features.

On the other hand, many smaller groups such as the ones that lived in the Amazonas and the Pampas regions had less centralised and more political autonomous ethnic features. These groups were organised into small bands, and their ways of living did not require to run systems capable of administering complex political activities other than the necessary tasks that allowed them to subsist independently.

At the time of contact, Europeans somehow understood these differences of ethnic power structures, as well as the strategic advantages that came with them. Besides the expected struggles, colonisers found it easier to subjugate and settle in areas where larger groups inhabited. The already established complex ethnic systems of more advanced groups allowed them to control the native population and extract resources rapidly through the help of local leaders. *Caciques* or *kurakas* provided, for at least a century or so, essential administrative and authority tasks over many small communities by generating the main tributes and supply of labour for the expansion of the colonial enterprise. A prominent envoy of the Spanish crown wrote (Saignes 1999, p. 65):

"The day we decide to count and tax all Indians individually, so that if they have paid they are no longer under the domination and authority of the caciques and under their orders, being free to go wherever they want, we will have taken away the restraint that holds them together in an orderly manner, for this is the way in which they survive and have survived before the Christians obtained these realms: if one could put this fact to the test for only one year, one would clearly see their destruction".

Recent studies have found a large variation of the economic activity undertaken in the colonial period within present-day Latin America's nations. As colonisation progressed, the relative stability that these areas offered well into the second century of colonisation may have enabled Europeans to develop some solid economic activities, as well as the main political and financial centres¹. Indeed, although the configuration of most - if not all- the ethnic homelands vanished as the colonial power entered in the continent, it seems that one important determinant for colonisers to settle and operate economically and politically was the degree of state capacity of ethnic groups.

¹It is important to note that while the literature points out that the areas with higher development gave way to the establishment of "extractive" economic activities -resulting in negative consequences for economic development over the long-run, recent evidence addressing the "reversal of fortune" hypothesis at the subnational level is somewhat less clear. Please see Bruhn and Gallego (2012) and Maloney and Valencia (2015) for two important contrasting views on such hypothesis.

3.4 Data and Methodology

3.4.1 Digitalisation of Murdock's historiographical maps

The present Chapter constructs a new dataset drawn from two main sources containing a series of Peter G. Murdock's historiographical publications of maps and ethnographic data portraying spatially the locations of ethnic homelands in Latin America as of the XVI century.

The first source comes from the *Outline of South American Cultures* (Murdock 1951). In this book, Murdock classified ethnic groups from Central and South America according to their main traditional ethnic homelands. Separated into 24 different maps to correspond the modern political divisions in this region, Murdock identified the spatial location of 216 ethnic homelands as of prior contact with Europeans. Each of these ethnic homelands are represented by the main ethnic group that historically inhabited them. For example, the map of what is today Colombia has a conglomeration of 24 ethnic homelands representing by a larger ethnic group such as Chibcha, Witoto, Catio, etc.

In addition to the ethnographic information on the spatial location of the ethnic homelands, Murdock also provided in this book ethnographic data summarising key ethnic characteristics. Amongst the data included was the institutional complexity of ethnic groups. As I will explain below, this information helps me expand the number of observations in my dataset by allowing me to assign the respective value of institutional complexity of some ethnic groups.

The *Ethnographic Bibliography of North America* (Murdock and O'Leary 1975) is my second source. In this material, Murdock and O'Leary provided the classification of ethnic groups in North America, including from the groups that historically inhabited the Mexican territory. As in the first source, authors outline the spatial location of ethnic homelands into different maps, as well as into a single hemisphere map. From this book, I use the ethnic homelands that fall exclusively within the territorial domains of Mexico, and append them to the rest of the maps so that the mainland of Latin American countries is fully covered in the empirical analysis.

Unlike the previous source, though, this second book does not provide a summary on the major ethnic characteristics. Nor did have a complete outline of the ethnic homelands in Mexico and some parts in Central America, in particular within the area along the Yucatan Peninsula and Guatemala -a culture area where Maya-related groups lived. Yet through this source another 45 ethnic homelands that historically inhabited by different

ethnic groups (e.g. such as Aztecs, Mixtecos, Tepehuanos, Seri, Huave, etc.) are added to the units collected from the previous Murdock's historiography material.

In order to empirically analyse the effects of the historical ethnic homelands in Latin America, I digitise these series of historiographical maps through the use of Geographical Information Systems (GIS). To do so, I project these maps on a geographical software by using the two-dimensional coordinates of grids as specified in my sources. This allows me to establish the geographical location of my maps on a flat surface so that the drawing of the ethnic homelands can be then performed consistently all along the digitalisation.

Figure 3.1 shows the full digitalisation of the ethnic homelands in Latin America. The black line represents the ethnic homelands boundaries prior contact with Europeans, whereas the red line depicts the contemporary administrative divisions of Latin American countries. A total of 375 ethnic homelands are created. Such increase in units is as a result of the intersection of the ethnic homelands with the current boundaries of countries making them fall across more than two neighbouring countries. For instance, the ethnic homeland of the Aymara group in South America is partitioned across three countries: Bolivia, Chile and Peru.

It is worth pointing out that the area along the coast between Venezuela and Brazil in South America is in blank. This area corresponds to the countries of Guyana, Suriname and French Guiana. I omit this land in order to be consistent with the analysis conducted in the previous chapter. In so doing, I only loose one ethnic group (Arawak). As such, another six ethnic groups that lived within the land of the Caribbean countries are also omitted. Finally, and as mentioned earlier, one can see too that the Yucatan Peninsula and Guatemala, as well as some land in the south of Honduras and El Salvador, were not included in the analysis due to the lack of ethnographic data². As a robustness check, though, I add in the empirical analysis only the omitted area that fall within Mexico and Guatemala by assigning the corresponding values of the most representative ethnic group that historically inhabited this region -the Maya group.

3.4.2 Constructing main predictor, outcomes and controls

To measure the institutional complexity of the ethnic groups that historically inhabited these ethnic homelands, I build on the pioneering anthropological data drawn from the

²None of the two sources that I used to digitise the ethnic homelands did not develop historiographical work of this particular area. Though, according to alternative sources such as GREG dataset (available at <http://www.icr.ethz.ch/data/other/greg>), I found that most of the missing area seems to be a region where mainly Maya-related groups live today.

Figure 3.1: Traditional Ethnic Homelands in Latin America



Murdock's Ethnographic Atlas (1967). From this source, I use the variable called *Jurisdictional Hierarchy beyond the Local Community*, which classifies the institutional complexity of ethnic groups as outlined before.

For this, I have to match the ethnic groups included in my new dataset to the ones classified in the Atlas. As a result, only around of one quarter of my observations receive a value of institutional complexity (135 observations out of 375). To overcome missing values in my data, I implement the following methodological strategy. I use the ethnographic data on institutional complexity of ethnic groups provided through the book: *Outline of South American Cultures (OSAC)*. The rule to follow is to input a value on an ethnic group from the OSAC source only if the Atlas does not provide it. So, in cases where the value on institutional complexity are recorded on both sources (Atlas/OSAC), I use the information as documented in the Atlas. Such strategy is supported by the fact that the correlation obtaining between these cases is very high (89%). This, thereby, enables me to expand up to two quarters the number of ethnic groups with a valid value on institutional complexity standing at 265. The analysis will use this sample of 265 observations for the main empirical investigation. Though, in the robustness section, and as conducted in the second Chapter of the present thesis, I assign a value of zero of Murdock's index -denoting stateless characteristics- to the ethnic groups that are still having missing values.

Nevertheless, the construction of our index measuring the institutional complexity of ethnic groups seems to have a measurement problem not found in my previous chapter. As noted before, the values of Murdock's *Jurisdictional Hierarchy beyond the Local Community* takes a categorical form ranging between 0 to 4. In our dataset none of our 265 observations report a value 3; only 2 take a value of 4; 28 a value of 2; 34 a value of 1; and 202 the lowest level, 0. To minimise measurement errors as result of outliers, I include all the groups reporting paramount chiefdom and state-related characteristics into level 2. Our index thus takes three categorical values: level 0 signifying stateless groups; level 1 representing groups with petty chiefdom features; and level 2 for groups with paramount chiefdom and state-related characteristics. To be consistent to the definition of measures in the present thesis, I call this variable as pre-colonial institutions (PCI), which takes ordinal values ranging from 0 to 2, where zero denotes groups with stateless features and 2 depicts groups with paramount chiefdom and state-related characteristics.

Figure 3.2 shows the variation of the measure of pre-colonial institutions of ethnic groups in Latin America as of the XVI century. The red colour denotes the ethnic homelands in which more advanced groups inhabited (paramount chiefdoms & state-related groups); the green colour represents the areas where petty chiefdoms lived; and the grey colour signifies the areas where less centralised and more political autonomous groups lived. The white areas represent the ethnic groups with no information of Murdock's

institutional complexity (as mentioned above, in our robustness check section below, I include these areas by assigning a value of zero of Murdock’s institutional complexity). As it can be seen, -all together- the variation is prominent, with some areas like the Gran Chaco and Amazonas in South America having the lowest value of pre-colonial institutions (level zero). Meanwhile, areas such as Central Mexico, the weastern and some parts in the north-eastern of South America concentrate ethnic groups with paramount chiefdom and state-related features (level two).

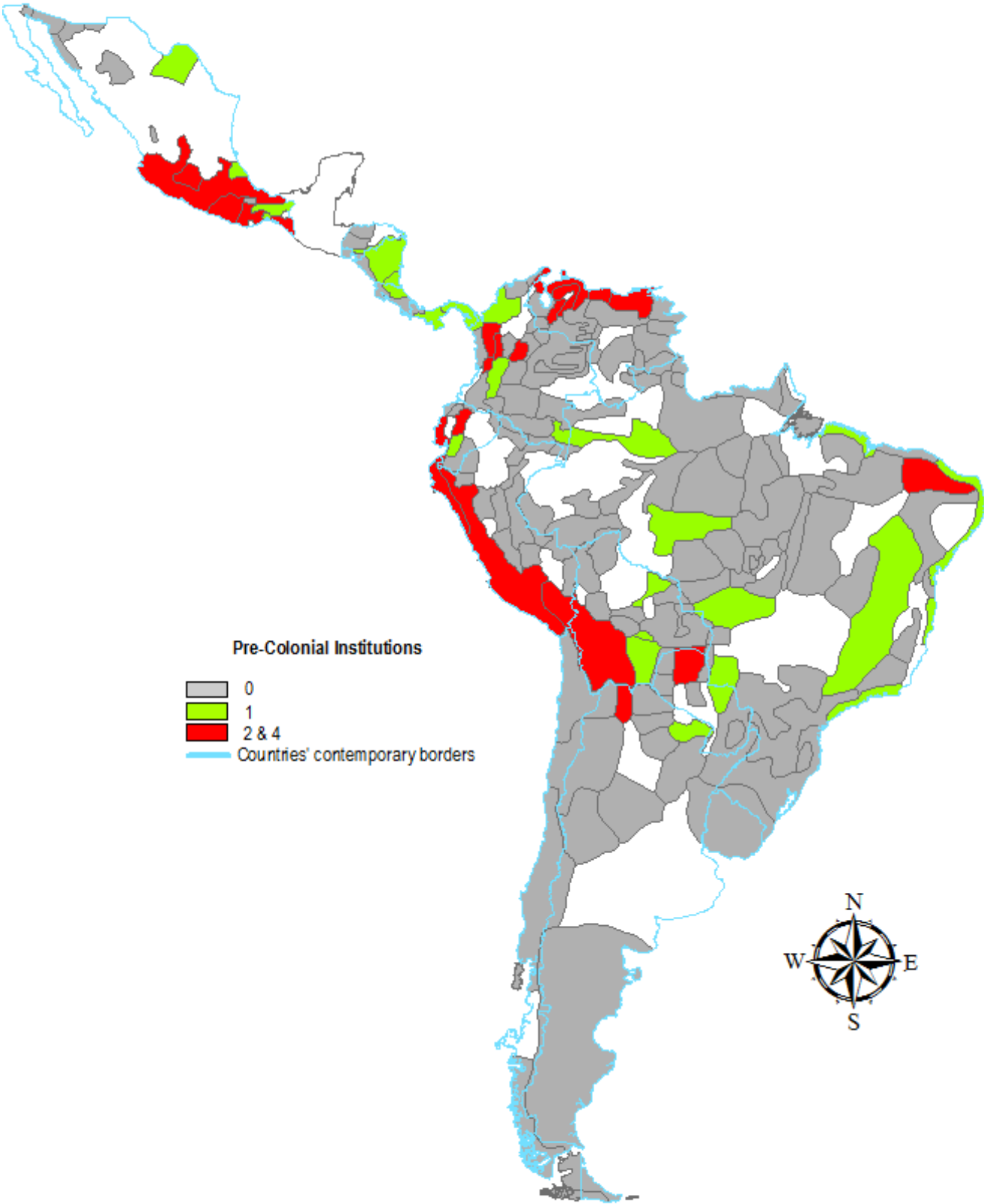
To measure the effects of pre-colonial institutions of ethnic groups on present-day economic development outcomes in Latin America, I build on recent work in using satellite light density data as a proxy for economic development (Henderson et al. 2012; Michalopoulos and Papaioannou 2013). The data is collected by the Defence Meteorological Satellite Program’s Operational Linescan System (DMSP-OLS) through satellite images of light density at night across the globe. This work is done in a systematic way throughout the year so that a series of images are gathered, scientifically cleaned from any disturbance (such as cloudy days, sunny spells, etc.), and averaged on a yearly basis. The data take discrete values ranging from 0 to 63, and is presented in raster files with a finer resolution of 30-second grid (similar as one squared kilometre). I intersect my ethnic homelands with the raster files on light density. Thereby, all the grids that fall within the ethnic homelands are averaged; as a result an overall level of economic activity at this novel level of analysis is captured. I use the data on light density from the period 2000-2013. This measure is my main development outcome throughout the empirical investigation.

Figure 3.3 illustrates the variation of economic activity at the level of ethnic homelands in Latin America -as measured by light density at night. As the thickness of the colour increases it would be an indication of areas with higher levels of economic development. At first sight, it seems that areas inhabited by more advanced groups such as the weastern of South America and Central America tend to reflect, on average, more light density at night, yet there are areas that were historically inhabited by less advanced groups with high levels of light density as well like in the southside of Latin America.

To attest the history of institutionally-advanced ethnic features in Latin America conditional on major fundamentals of nations’ development, I proceed as follows. I construct a large array of controls in a similar manner as my light density data. First, I am particularly interested in controlling for the following geographic factors:

- Land quality: This is a geospatial index that examines the land suitability for agriculture through three major geographic components: croplands, climate conditions and soil characteristics (Ramankutty et al. 2002). It takes values between 0 and 1,

Figure 3.2: Pre-Colonial Institutions within Ethnic Homelands in Latin America



with higher values denoting more fertile land. I compute this index by averaging all the grids that fall within each ethnic homeland.

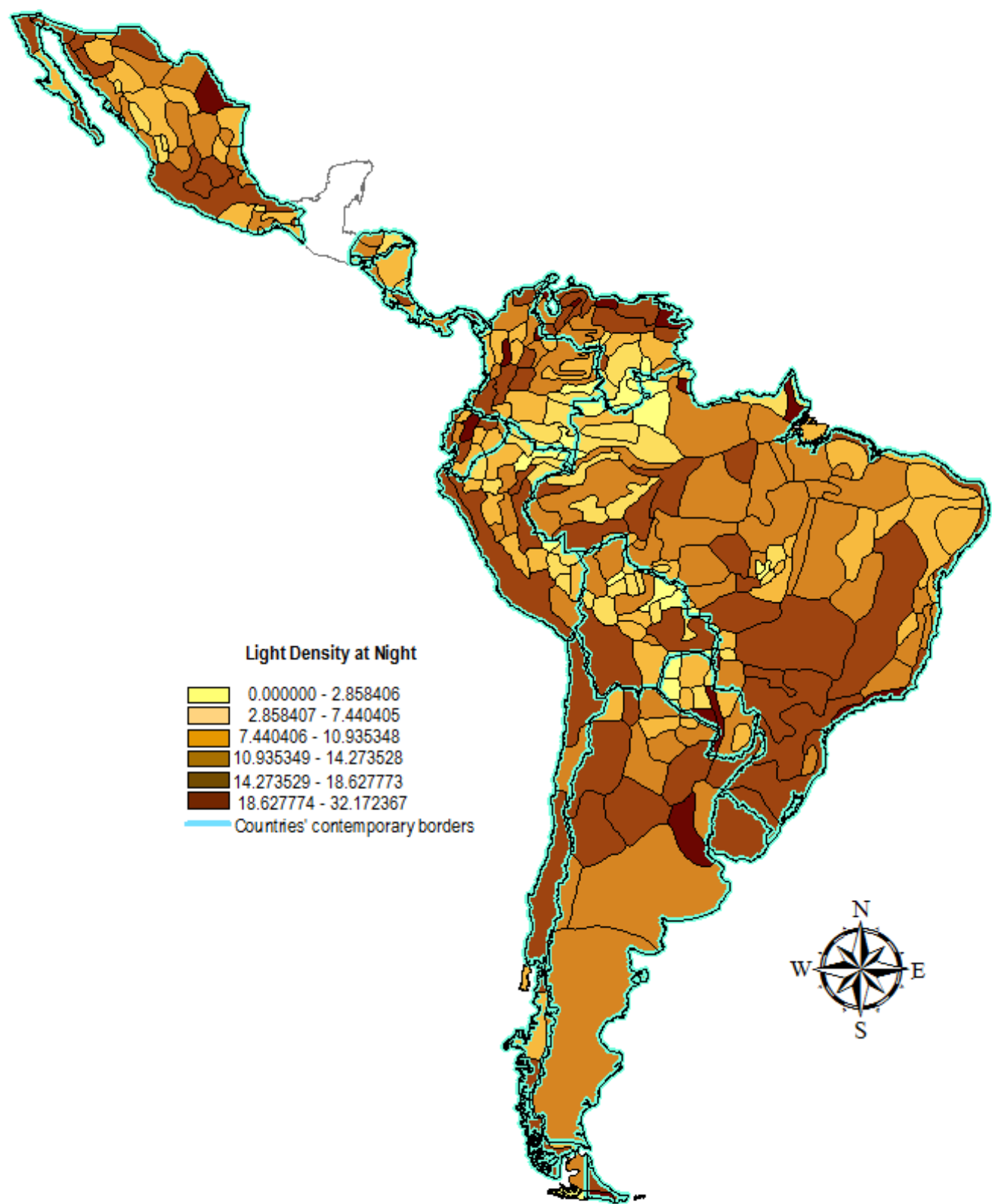
- Elevation: By computing the average of all the grids in each ethnic homeland, I obtain the average elevation at the required level of analysis.
- Temperature: This is the mean yearly temperature data over the period 1950-2000. I average all the grids falling within my unit of analysis.
- Precipitation: This is the mean yearly precipitation data (mm) over the period 1950-2000. I compute the data as my temperature variable above.
- Malaria: This is an index capturing the stability of malaria transmission across the world (Kiszewski et al. 2004). I calculate the average of the grids for each ethnic homeland.
- Latitude: I compute the absolute latitude in each ethnic homeland.

Second, it is still possible that the association between my measure of pre-colonial institutions and contemporary economic development is affected due to other omitted factors related to the demographical or technological success of nations driving the paths of development in the long-run. As such, I also control for the following determinants of development:

- Oil & Gas: Using geospatial data drawn from the project called Peace Research Institute Oslo (PRIO), I create an indicator equals one if there is an oil or gas field within an ethnic homeland, and zero otherwise.
- Population Density in Logs: Using the population digital data (as of 2000) from the United Nations Environment Program, I calculate the population density for each unit (in logs) by taking the average of all the grids as before.
- Nearest Distance to Capital: For each country, I calculate the nearest distance of ethnic homelands to the capitals. For example, using the centroid of each ethnic homeland within the Mexican territory, I compute their nearest distance (in kilometres) to Mexico City.

The next step is to implement a method that aims to account for alternative mechanisms mediating the distribution of economic development in the long-run. As noted before, colonial factors served as a potential exogenous variation of nations' economic success throughout time. Colonisers transformed the systems that they encountered at their

Figure 3.3: Light Density at Night within Ethnic Homelands in Latin America



arrival in the New World, affecting since then the direction of development of today's Latin American countries. Yet such transformation took place based on -but not only- the different ethnic institutional advancements established prior colonisation. As colonisers settled in areas characterised from having more advanced ethnic groups, their main economic activities were more likely to be established, and for consequence, the political and financial institutions that held them. As already documented in the literature, this friction may have affected, by and large, the development of regions in Latin America giving rise to a higher economic outcome over time in places that guaranteed during the colonial period more solid European activities as a whole.

Thus, I construct a crude measure of economic activities implemented during the colonial period within my ethnic homelands. For that, I rely on my state-level data that was prepared for the previous chapter. From this data, I use the following historical indicators measuring the within-countries variation of colonial activities (Bruhn and Gallego 2012):

- Other colonial activities: An indicator that takes a value of one if a subnational state in the Americas developed economic activities related to agriculture during the colonial period, and zero otherwise.
- No colonial activities: An indicator that takes a value of one if a subnational state in the Americas did not develop economic activities during the colonial period, and zero otherwise.
- Plantations: This is also an indicator taking a value of one if plantations were recorded during the colonial period in a subnational state across the Americas, and zero otherwise.
- Mining: This indicator measures the presence of mining activities in the colonial period at the subnational state level by taking a value of one if such economic colonial activity was present and zero otherwise.

To capture the variation of colonial activities at the ethnic homeland level, I intersect our map of ethnic homelands with a map containing all the contemporary boundaries of subnational states in Latin America. As a result, the ethnic homelands fall within the boundaries of one or more subnational states. For each historical indicator, I then simply compute the average of the respective values of the subnational states falling in each ethnic homeland. For instance, Figure 3.4 shows the intersection of the ethnic homeland of the Catio group -an ethnic group reporting centralised institutional features. As it can be seen, the area of the ethnic homeland of this group is mainly falling across four subnational states in Colombia: Cordoba, Antioquia, Choco and Risalda. By using the

information of the indicator on other colonial activities from my state-level data, we know that the first two subnational states recorded other economic activities during the colonial period (taking a value of 1), whereas the Choco and Risalda states did not (taking a value of zero). So, the value of other colonial activities for the ethnic homeland of the Catio group is the average of these four dummies. I use the same strategy to compute the remaining indicators on plantations, mining and no colonial activities. Thus, four separate variables are created measuring the type of colonial activities that may have been taken place within the fringes of the ethnic homelands³.

In addition to the above measures on colonial activities, I am also in the position to exploit other relevant historical information along with present-day statistics from my state-level data. As regards the historical information, I am particularly interested in measuring pre-colonial population density. This information would be useful to control for pre-colonial development other than the institutional advancements of ethnic groups. Moreover, I am also interested in using statistics measuring the today's share of Amerindian population. This information will be useful to control for the discriminatory effects transmitted via the Amerindian population, picking up as a result a biased downward effect on the coefficient for pre-colonial institutions.

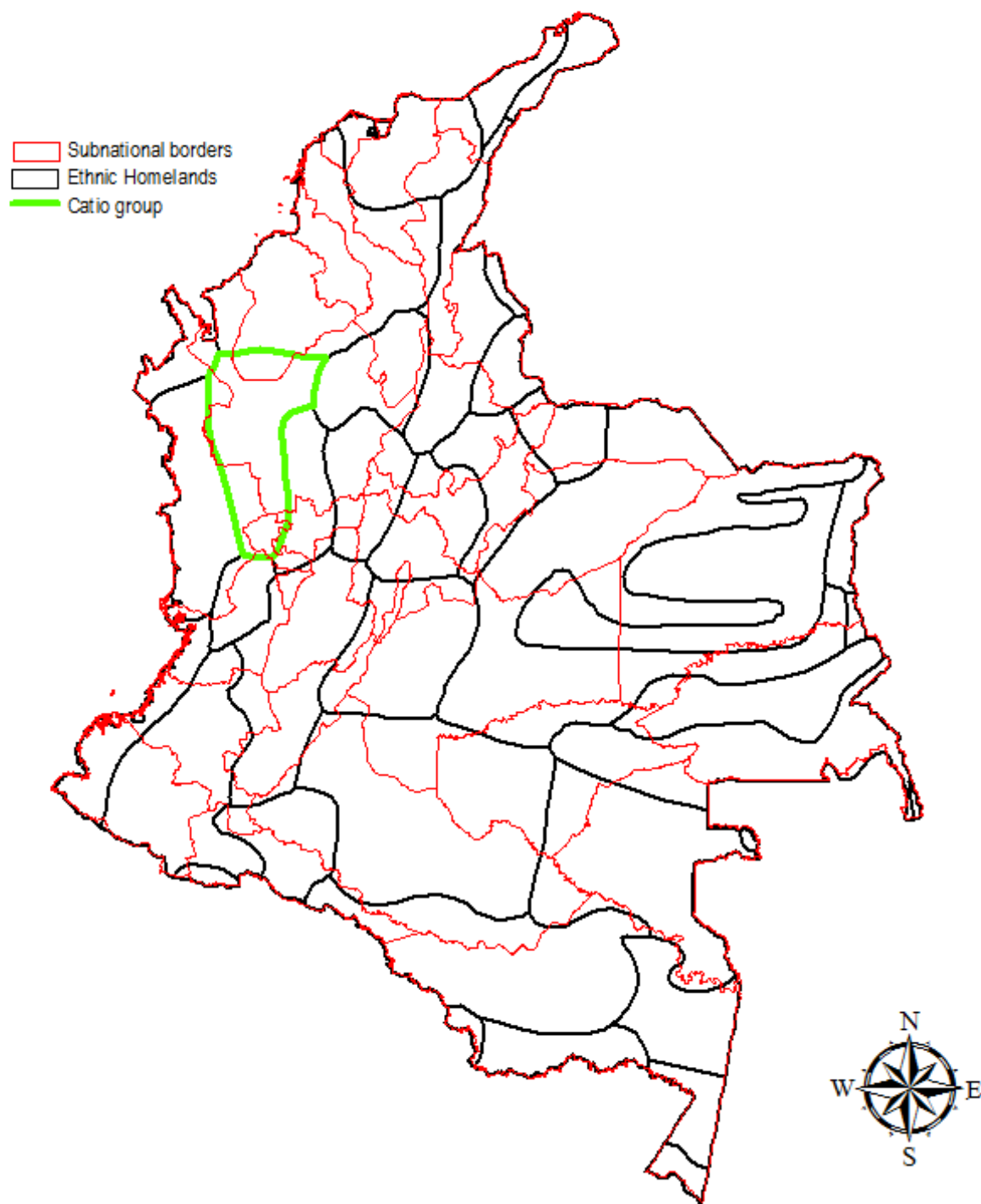
It is important to note that, as in the indicators on colonial activities, these two additional variables are computed by calculating the average of the respective values of the subnational states falling in the ethnic homelands. For example, using the same ethnic group as above, the value on pre-colonial population density for the Catio group in Colombia will be equal to average of the values on pre-colonial population density corresponding to the subnational states of Cordoba, Antioquia, Choco and Risalda.

3.4.3 Descriptive statistics

As Panel A of Table 3.1 shows, the number of ethnic homelands within the contemporary countries in Latin America vary, as well as the territorial size of them. The amount of ethnic homelands in nine countries ranges between 12 to 46 -with the exception of Brazil with the largest number in the dataset at 102. The other six countries have relatively less ethnic homelands from only 1 (El Salvador and Uruguay) to 6 (Honduras). The

³As an alternative strategy, I created indicators assigning a value of one if an ethnic homeland is intersected by a subnational state reporting these colonial activities. For instance, if an ethnic homeland falls within three subnational states and only one reports these colonial activities, this unit takes a value of one. Though, such strategy did not change the main results of the present research.

Figure 3.4: Intersection between Ethnic Homelands and Subnational States (Colombia)



territorial domains of these ethnic homelands also oscillate plenty, from the biggest sizes corresponding to the groups like Tupi, Incas, Aztecs, Araucanias, Charruas, with a land area ranging as much as 616,000 sq. km., to the smallest ethnic homelands with a land area as lower as 61 sq. km (Encabellado group).

The summary statistics of our main variables is in Panel B of Table 3.1. The continuous values of the dependent variable on light density range between 0 and 32, yet its distribution is fairly stable as depicted by the value of the standard deviation. In Panel B, I also report the summary statistics on our main independent variables. As it can be seen, only 265 observations received a valid value of institutional complexity. Finally, I am also reporting in the last rows of Panel B all the array of geographical, location and historical controls. Please refer to Table B.1 of Appendix B for specific details on the main sources and description of variables.

3.5 Empirical strategy and main results

3.5.1 Econometric model

To examine the long-run effects of Pre-Colonial Institutions on the basis of the traditional ethnic homelands, I use the following main specification:

$$Y_{e,c} = \beta PCI_{e,c} + \theta popden_{e,c} + \ell Amepop_{e,c} + \lambda \mathbf{X}_{e,c} + \alpha_c + \varepsilon_{e,c} \quad (1)$$

In equation 1 subscript e indicates ethnic group and subscript c signifies countries. $Y_{e,c}$ represents our measure of light density at night as a proxy for economic activity at the level of ethnic homeland. $PCI_{e,c}$ is our categorical measure of pre-colonial institutions. $popden_{e,c}$ is controlling for population density in modern days and $Amepop_{e,c}$ for today's share of Amerindian population. $\mathbf{X}_{e,c}$ is our main vector of controls for geographic and location factors. Country fixed effects are also included $-\alpha_c$. Our standard errors are clustered at both country and ethnic homeland levels following the approach developed by Cameron, Gelbach, and Miller (2011); and applied in Michalopoulos and Papaioannou (2013).

Table 3.1: Summary statistics

Panel A					
Country	Nun. Ethnic homelands	Min. land area (sq. km.)	Max. land area (sq. km.)		
Argentina	27	1260.883	1291940		
Bolivia	35	582.2659	293066		
Brazil	102	294.8566	616320		
Chile	12	7562.352	333230		
Colombia	39	517.9127	150851		
Costa Rica	3	13678.45	20544.39		
Ecuador	17	277.481	41445.74		
El Salvador	1	7780.768	7780.768		
Honduras	6	766.069	34009.07		
Mexico	46	993.1024	306319		
Nicaragua	4	716.8536	102069		
Panama	3	10941.22	33622.36		
Paraguay	15	2010.41	96656.65		
Peru	34	61.64779	379921		
Uruguay	1	250640	250640		
Venezuela	30	667.3262	88899.95		
TOTAL	375				
Panel B					
Variables	Observations	Mean	Std. Dev.	Min.	Max.
<i>Main dependent variable:</i>					
Average Light Density at Night	375	10.1049	5.849169	0	32.17237
<i>Main independent variables:</i>					
Pre-Colonial Institutions	265	0.3509434	0.6752168	0	2
<i>Controls:</i>					
Population density in modern days (logs)	375	2.037937	2.483619	-4.60517	6.604855
Share of Amerindian population	375	0.1233403	0.1408891	0.0010034	0.7845646
Malaria	375	2.509646	1.837855	0	5.666667
Absolute latitude	375	14.22596	11.32516	0.1166143	54.80722
Land Suitability Index	375	0.3655606	0.2851374	0	0.998
Elevation	375	639.8869	763.7535	5.083333	4415.238
Mean Yearly Precipitation in mm	375	144.6934	76.68576	5.354762	517.9863
Mean Yearly Temperature in Celsius	375	21.87064	5.45098	0.7953707	28.01937
Distance to capitals (in km.)	375	392386.7	392984.5	0	1409249
Oil & Gas Dummy	375	0.3013333	0.4594502	0	1
Pre-colonial population density (logs)	374	0.0793835	1.416842	-5.120699	3.420635
Other economic activities	372	0.5867498	0.3433686	0	1
Plantation	372	0.0764325	0.1893317	0	1
Mining	372	0.1625331	0.2722902	0	1
No colonial activities	372	0.1742845	0.2776389	0	1

3.5.2 Empirical results

Table 3.2 reports the association between the legacy of pre-colonial institutions and contemporary economic development at the ethnic homeland level in Latin America. Our specifications use as dependent variable light density at night to proxy for economic activity at this level of analysis. The large array of controls will be included gradually so that we systematically inspect the response of our main independent variable measuring the legacy of pre-colonial institutions. Since our development outcome takes a continuous form, OLS estimators are used throughout the specifications.

We start by including in specification (1) of Table 3.2 our variable on pre-colonial institutions; our control for population density in 2000s; plus country-fixed effects⁴. The inclusion of fixed effects allows me not only to exploit within-country variation but also to control for unobservable country-specific characteristics. As noted in my previous Chapter, I am particular interested in controlling for some colonial and post-colonial institutional factors affecting the variation of my main development outcome. Moreover, due to the fact that some of my units of analysis are partitioned in more than two countries, plus the possibility that intra-related ethnic characteristics may be correlated within major ethnolinguistic families, I am clustering my standard errors at both country and ethnic levels. As it stands, the results support the main thesis of the present research. The coefficient capturing the legacy of pre-colonial institutions is positive and significant at 5% level.

In column (2), I augment my specification by including the variable on the share of Amerindian population in modern days. As expected the sign of the coefficient is negative but it cannot explain the variation on our outcome at all. More importantly, by adding this control our coefficient for pre-colonial institutions increases in magnitude and significance levels. This may suggest a biased downward effect on the coefficient for pre-colonial institutions caused by the discriminatory effects transmitted through the Amerindian population -as shown in my previous chapter.

I finally add in column (3) our large array of geographic, locational and natural resources controls. Recent findings in Maloney and Valencia (2015) suggest that the evolution of the pre-colonial, colonial and post-colonial settlements have been affected by locational fundamentals -along with agglomeration effects. For instance, ethnic groups prior colonisation tended to be located in more suitable areas; Europeans followed suit during the colonial period; so did the societies in modern days. If that is the case then all these factors but the legacy of institutional complexity are the main drivers of the variation

⁴Please note that as in Chapter 2, population density may be regarded as a "bad control". Thus, interpretation of results using this variable should be cautiously.

Table 3.2: Baseline results

Dependent variable: Light density at night (2000-2013)			
Variables:	(1)	(2)	(3)
Pre-Colonial Institutions	0.702** [0.326]	0.777*** [0.288]	0.634*** [0.187]
Population density (2000s)	1.618*** [0.0839]	1.562*** [0.0908]	1.433*** [0.233]
Share of Ame. Pop.		-2.988 [2.244]	-2.084 [2.285]
Absolute latitude			0.0731 [0.0568]
Precipitation			-0.630 [0.617]
Temperature			1.679 [1.362]
Elevation			0.477* [0.256]
Malaria			-0.327 [0.337]
Land Suitability Index			-0.394 [0.568]
Distance to capitals (km.)			6.07e-07 [2.10e-06]
Oil & Gas Dummy			1.853** [0.737]
Country Fixed Effects	YES	YES	YES
Observations	265	265	265
Adjusted R-squared	0.512	0.514	0.553

Notes: Cluster standard errors at country and ethnic levels are in brackets. *** p<0.01, ** p<0.05, * p<0.1

of our economic development outcome within the geographical areas under study. All in all, the results reassure our main hypothesis: pre-colonial institutions relate to better levels of economic development in ethnic homelands historically inhabited by more advanced ethnic groups in Latin America.

While our results already control for unobservable country-specific characteristics and a large array of geographical and locational fundamentals, other important historical factors may be the main source of variation of our development outcome. As noted above, Europeans settled in areas where more advanced ethnic groups lived as a result of key locational fundamentals and agglomeration effects (Maloney and Valencia 2015). This in turn gave way to different types of colonial activities all along Latin America triggering a potential exogenous variation of nations' economic success throughout time. If all these historical factors have an effect on the economic activity undertaken within our units of analysis, the causal channel that I hypothesize between pre-colonial institutions and economic development would then be through these historical factors.

To test the above alternative channels, in column (1) of Table 3.3, I add our measure of pre-colonial population density as a relevant control for pre-colonial development other than the institutional complexity of ethnic groups. The specification includes our variable of interest, along with the comprehensive set of geographical and locational controls, plus country fixed effects. The results strongly support our thesis. The coefficient for pre-colonial institutions remain unchanged. Meanwhile, the coefficient for pre-colonial population density is not statistically significant different from zero. Though, we obtain a negative association. This may go in line with our previous analysis and the main results in the existing literature pointing out towards what is known as the "reversal of fortune" (Acemoglu et al. 2002; Bruhn and Gallego 2012).

In column (2), I augment the specification by adding the set of indicators for the different colonial activities leaving the indicator for no colonial activities as our reference category. Once again, our coefficient for pre-colonial institutions remain positive and very significant suggesting that the type of colonial activity is not mediating the variation of economic development within our units of analysis. On the other hand, in terms of the effects of colonial activities, all the indicators show a negative association with our main development outcome but with insignificant effects⁵.

Finally, I am also in the position to implement an empirical strategy to examine the accuracy of our measure on pre-colonial institutions. First, our categorical variable on pre-colonial institutions is transformed into an indicator taking the value of one to ethnic

⁵Please note that as in Chapter 2, colonial activities may be regarded as a "bad control". Thus, interpretation of results using these variables should be cautiously.

Table 3.3: Pre-Colonial Development and Colonial Activities

Dependent variable: Light density at night (2000-2013)		
Variables:	(1)	(2)
Pre-Colonial Institutions	0.666*** [0.198]	0.651*** [0.190]
Pre-colonial population density	-0.303 [0.326]	-0.262 [0.400]
Other Colonial Activities		-1.223 [0.966]
Plantation activities		-1.582 [1.124]
Mining activities		-0.0358 [1.458]
<i>Controls for:</i>		
Population Density (2000s)	YES	YES
Share of Ame. Pop.	YES	YES
Geography, location, and natural resources	YES	YES
Country Fixed Effects	YES	YES
Observations	264	262
Adjusted R-squared	0.555	0.557
Notes: Cluster standard errors at country and ethnic levels are in brackets. *** p<0.01, ** p<0.05, * p<0.1		

groups reporting paramount chiefdom and state-related characteristics (values 1 and 2 of Murdock's index), and zero to groups reflecting stateless features (level zero). The aim of this aggregation is to reduce measurement errors in Murdock's categorical index.

In Table 3.4, column (1) reports results using the dummy form of our variable on pre-colonial institutions. Of course, I am also including all the array of controls as in the previous specifications from Table 3.3. For abbreviating reasons, I am not reporting the coefficients of this set of controls. The results once again support the main thesis of the present chapter. The coefficient on pre-colonial institutions increases in magnitude and remains statistically significant, though at 5% level.

Second, I create three separate indicators measuring the different levels of institutional complexities of ethnic groups: i) stateless groups, ii) petty chiefdoms, and iii) paramount chiefdoms and state-related groups. Before reporting the econometric results, I have graphically visualised in Figure 3.5 these indicators against our main development outcome. The y-axis shows the mean of light density, whereas along the x-axis these indicators are piled up side-by-side. From this data visualisation, it is clear the larger differences of economic

Figure 3.5: Categories of Pre-Colonial Institutions and Light Density at Night

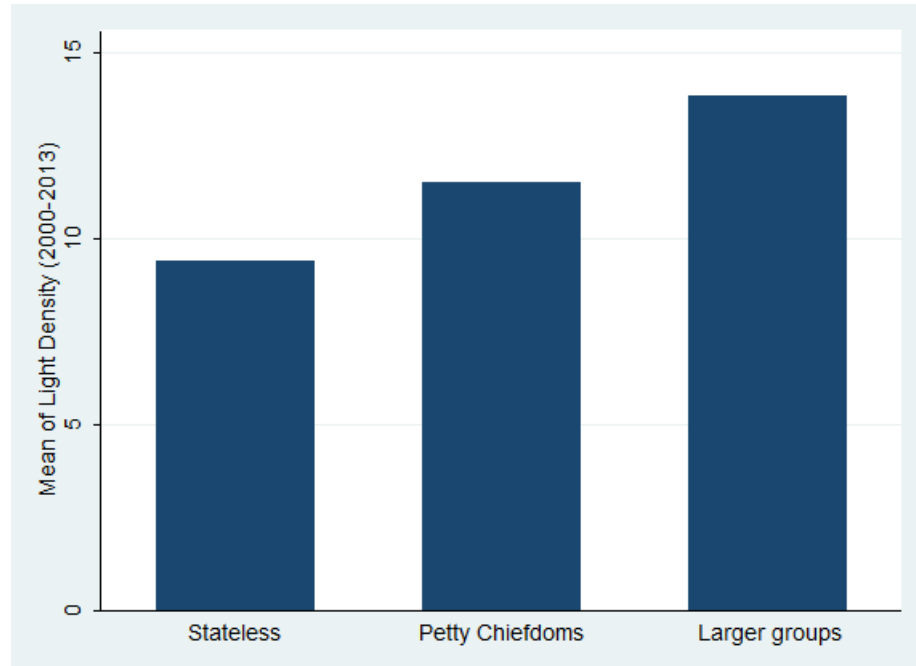


Table 3.4: Pre-Colonial Institutions (dummy approach)

Dependent variable: Light density at night (2000-2013)		
Variables:	(1)	(2)
Pre-Colonial Institutions (dummy)	0.958** [0.417]	
Petty Chiefdoms		0.657 [0.621]
Larger groups		1.299*** [0.256]
<i>Controls for:</i>		
Population Density (2000s)	YES	YES
Pre-colonial population density	YES	YES
Colonial activities	YES	YES
Share of Ame. Pop.	YES	YES
Geography, location, and natural resources	YES	YES
Country Fixed Effects	YES	YES
Observations	264	262
Adjusted R-squared	0.559	0.557

Notes: Cluster standard errors at country and ethnic levels are in brackets. *** p<0.01, ** p<0.05, * p<0.1

activity within the ethnic homelands depending on the legacy of institutional complexity of ethnic groups. Areas that were historically inhabited by more advanced groups tend to reflect on average 30% more light density at night as compared to the areas where groups with stateless characteristics lived, and around of 11% higher with respect to the petty chiefdom groups.

In columns (2) of Table 3.4, I report the specifications to explore these individual effects of institutional complexity of ethnic groups, conditional on observable and unobservable country-specific characteristics. The specification uses the two indicators for petty chiefdoms and larger groups all together, leaving the indicator for stateless groups as our reference category. The results remain in favour of our thesis. The coefficient on larger ethnic groups is highly significant and very large in magnitude.

3.5.3 Robustness checks

The robustness of the results is tested through various alternative approaches. First, as one could see from Figure 3.1, the Yucatan Peninsula and Guatemala were not included in the analysis due to the lack of ethnographic data. This is a relevant area not only because of the large size of land (225,000 sq. km. approx.) but also due to the fact that it is a culture area where Maya-related groups lived (Quiches, Chori, Cakchique, etc.). I include this land as part of the Maya group taking thus a value of 1 of Murdock's institutional complexity. In Table 3.5, specification in column (1) uses our benchmark dataset with 265 ethnic homelands plus the homeland of the Maya group. Light density is our main outcome; our main independent variable is our categorical variable of pre-colonial institutions; and all the controls are also included as of specifications of Table 3.4. The main independent variable is positive and significant confirming again the thesis of the present research.

Second, one may also want to think of the Maya group as one of the most advanced pre-colonial civilisations in Central Mexico. I then assign a value of 2 of Murdock's institutional complexity to the area where Maya-related groups lived -in line with the values of the larger groups in this region (e.g. Aztecs, Tarascos, Zapotecs). Column 2 of Table 3.5 shows the results using this approach. As it can be seen, our main coefficient dropped by a third of its size, yet the correct sign is still positive and statistically indistinguishable from zero at 1% level.

Third, I expand the size of observations in the dataset in an effort to cover all the ethnic homelands digitised in Latin America. As one can see from Figure 3.2, there are some ethnic groups that were left without information of Murdock's institutional complexity

Table 3.5: Robustness checks

Dependent variable: Light density at night (2000-2013)			
Variables:	(1)	(2)	(3)
Pre-Colonial Institutions	0.653*** [0.189]	0.456*** [0.177]	0.501** [0.235]
<i>Controls for:</i>			
Population Density (2000s)	YES	YES	YES
Pre-colonial population density	YES	YES	YES
Colonial activities	YES	YES	YES
Share of Ame. Pop.	YES	YES	YES
Geography, location, and natural resources	YES	YES	YES
Country Fixed Effects	YES	YES	YES
Observations	263	263	372
Adjusted R-squared	0.557	0.556	0.551
Notes: Cluster standard errors at country and ethnic levels are in brackets. *** p<0.01, ** p<0.05, * p<0.1			

due to the lack of ethnographical data. These groups come from mainly the area of northern Mexico and some parts in South America. I assign the lowest value of 0 of Murdock's institutional complexity to these groups, assuming that smaller groups would be more likely to have been overlooked by anthropologists. This conjecture is confirmed by the fact that ethnographical data are available for all the larger groups, plus the fact that most of the ethnic groups with no institutional complexity information come from the northern of Mexico -an area characterised to have had societies with more hunter-gathering and chiefdom features. This, thereby, allows me to include all the ethnic homelands digitised (375 in total). In Table 3.5, the specification in column (3) uses this full dataset covering all the ethnic homelands. Light density is our dependent variable, and as our main independent variable I am using our categorical variable of pre-colonial institutions, along with all the controls as of specifications of Table 3.4. The results fully reassure our previous findings.

3.6 Conclusions

The present Chapter studied the long-run effects of pre-colonial institutions on contemporary development in Latin America by relying on a novel dataset pinpointing the geospatial

location of the traditional ethnic homelands. By using ethnic homelands as the main unit of analysis, I was able to i) consider a different measure of economic development (satellite light density data); ii) relax the main assumptions of the central hypothesis of the present thesis; iii) compute our index of pre-colonial institutions on the basis of the single or more representative ethnic group that inhabited these units; and iv) mitigate -once again- identification problems arising from cultural changes, historical shocks, or colonial and post-colonial institutions with the use of country-specific fixed effects. The main results showed that the effects of pre-colonial institutions relate to a higher light density –as a proxy for economic activity- in ethnic homelands where more advanced ethnic groups lived. These results held even after controlling for major observable and unobservable country-specific characteristics, as well as with various methodological alternatives, samplings, and econometric permutations.

Chapter 4

On the Mechanisms of Long-Run Effects of Pre-Colonial Institutions in Latin America

Abstract

This Chapter studies a mechanism linking the persistence of pre-colonial institutions in Latin America over the long-run: Colonial and post-colonial strategies along with the ethnic political capacity worked in tandem allowing larger Amerindian groups to "support" the new political systems in ways that would benefit their respective ethnic groups -and the rest of the population. This mechanism, in turn, may have allowed the effects of pre-colonial institutions to influence socioeconomic development outcomes up to today. To shed lights on this mechanism, I combine the index of pre-colonial institutions prepared for the second chapter of the present thesis with individual-level survey data on people's attitudes. By controlling for key observable and unobservable country-specific characteristics, the main empirical results show that areas with a history of more advanced pre-colonial institutions increase the probability of individuals supporting present-day political institutions.

4.1 Introduction

The literature studying the relationship between institutional factors and economic development in the long-run has progressed significantly over the last years. Within this area of study, economists have placed special attention to the effects of pre-colonial institutions on present-day economic development outcomes

The fundamental contribution of this literature is that it distinguishes the existence of a pre-colonial ethnic structure below the national system that, to a certain extent, has

been influencing socioeconomic development outcomes in the long-run. As one of the most relevant ethnic characteristics, the institutional complexity that ethnic groups had at the time of the European arrival stands as one of the most powerful predictors of today's socioeconomic development outcomes.

Most of the contributions within this literature have concentrated on studying the actual relationship between pre-colonial institutions and economic development, but very little has been done to explain empirically the mechanisms through which this link has taken place. This research aims to contribute on this by looking at a mechanism already progressed in the literature related to the colonial and post-colonial strategies based on an indirect rule and the political capacity of more advanced ethnic groups to better process the new political system (Acemoglu et al. 2013; Gennaioli and Rainer 2007).

In the context of Latin America, it seems that the colonial and post-colonial strategies based on an indirect rule, along with the ethnic political capacity worked in tandem allowing larger Amerindian groups to "support" the new political systems in ways that would benefit their respective ethnic groups. As it will be outlined in the historical context, these strategies were implemented due to -but not only- i) the unfeasibility of the colonial power to control the large territories in Latin America, ii) the creation of autonomous Amerindian communities during the colonial period; and iii) the political autonomy granted to some Amerindian communities by the post-colonial political systems. All this allowed the Amerindian groups in Latin America to preserve some of their traditional forms of political organisations and, more important, to process on the basis of their political capacity the new systems introduced throughout the colonial and post-colonial periods in ways that would benefit their respective ethnic group and, therefore, their local communities, plus the rest of the population at large.

The positive interaction between the political capacity of ethnic groups and the new political systems may be understood as an alternative explanation for shedding lights on the mechanism by which the long-run effects of pre-colonial factors have affected socioeconomic development outcomes until today. I hypothesise that areas with a history of more advanced pre-colonial institutions are associated to a higher support of individuals for the contemporary political institutions.

To test the above hypothesis, this research combines the state-level data prepared for the second chapter of the present thesis and survey data on people's attitudes drawn from the Latinobarometer. From my state-level data, I use the index measuring the levels of pre-colonial institutions in each subnational state of Latin America. On the other hand, with the information available in the survey data, I construct four alternative outcomes measuring the support of individuals for contemporary political institutions (states, electoral bodies, national congress, and public administration).

First, I estimate an overall long-run effect by using all the individuals in the survey regardless their ethnic origins. The main empirical results show that areas with a history of more advanced pre-colonial institutions increases the probability for individuals supporting national institutions. Second, since the legacy of pre-colonial institutions would not be the same across all the different ethnic groups, I explore these long-run effects on individuals' attitudes by interacting our index of pre-colonial institutions with the information of the survey on the different ethnic identities of individuals. The results show that the effect of pre-colonial institutions vary depending on the different ethnic identities of individuals. This may offer evidence of some positive externalities coming from the Amerindian communities on other groups of the population.

The rest of the Chapter is structured as follows. Section 4.2 presents the main links of the present research with the literature addressing the economic differences in the long-run. In section 4.3 we provide a historical context placing special efforts to make the case for the mechanism under study. Section 4.4 outlines the data and methodology. Section 4.5 presents the main empirical results. Section 4.6 concludes.

4.2 Related literature

This research is linked to the growing body of comparative development studies addressing the long-run effects of ethnic-specific characteristics in the developing world (Gennaioli and Rainer 2007; Nunn 2008; Nunn and Wantchekon 2011; Michalopoulos and Papaioannou 2013; Fenske 2014; Alsan 2015).

As a starting point, the present research advances over the studies analysing the effects of pre-colonial institutions on contemporary development. These studies place great emphasis on the neglected existence of the dual structure -national versus ethnic- of developing countries downplaying the role of the central governments as the main political entities in shaping the economic and political factors. Relying on the large variation of institutional complexities that ethnic groups had prior colonisation, these studies found a positive relationship between pre-colonial institutions and different measures of socioeconomic development outcomes. Yet despite that the main conjectures of these studies were consolidated on the basis of key historical mechanisms, these latter remain to be empirically explored. By exploiting the variation of individuals' attitudes towards the political system through the added information on the levels of pre-colonial institutions, the present study explores, specifically, the mechanism related to the ethnic political capacity.

Second, this empirical investigation is also linked to the very few and recent studies analysing the actual mechanisms of ethnic-specific characteristics (Michalopoulos and Pa-

paioannou 2013). Perhaps the closer study -although more contextualised than empirical- to the present examination is the one by Acemoglu et al. (2013). They introduced the indirect rule set up by the colonial power in Africa as a potential mechanism to explain the positive relationship between pre-colonial institutions and socioeconomic development outcomes. They argue that after independence such indirect rule was more likely to be removed in areas with more advanced pre-colonial institutions allowing as a result the post-colonial government to monopolise the political power and thus bring social and economic stability for the whole nascent country. On the other hand, in the areas where stateless groups predominated, the colonial indirect rule endured rather than vanishing throughout the post-colonial period making difficult for the new political government to spread its political monopoly.

Third, the study also provides evidence on the state capacity developed by more advanced ethnic groups in pre-modern times. Diamond (1997) canonical hypothesis, for example, emphasises the capacity of larger groups to operate a more complex and multi-layer bureaucratic organisation, capable enough to collect resources and redistribute them in a form of a public good, for instance. The main findings of the present study show that, indeed, more advanced groups in Latin America -as captured by the addition of the levels of pre-colonial institutions- relate to an increase of population supporting the political institutions, suggesting an interlink between the past and present of the political capacity of ethnic groups.

An influential literature on the persistence of institutional factors is also linked to the line of research of the present chapter. With the theoretical basis laid down by North (1990), this literature has put renewed interest in providing empirical evidence on the deep historical roots of contemporary political and economic development of developing countries (Acemoglu et al. 2001; Bockstette et al. 2002; Banerjee and Iyer 2005; Angeles 2011; Dell 2010). My study shows that the pre-colonial institutional advancements of ethnic groups matter in the subsequent development of a stronger political system in the long-run.

Finally, the research contributes to the recent studies undertaken in other fields outside economics such as political science. These studies have focused on the emergent and relevant role of the Amerindian population into the contemporary political apparatus (Deborah 1998; Deborah 2005; Van Cott 2010). Although the main focus of these studies is on the electoral political participation of ethnic groups in Latin America, the analytical significance of this literature is the recognition of the capacity of ethnic groups in shaping the political landscape in Latin America. Yet these studies have not been able to explain why is that some ethnic cleavages tend to be more political salient than others. The present analysis provides an empirical explanation on this by introducing the legacy of

state capacity of ethnic groups as an important factor for some ethnic cleavages in shaping the political context of Latin American countries.

4.3 Historical context

Historically, the Amerindian groups have been at the front of major political changes in Latin America, yet their degree of influence in the new political context has depended on their political capacity along with the type of relationship with the colonial and post-colonial powers based -principally- on an indirect rule.

As has been noted before, although colonisation caused a radical transformation of the structure of the Amerindian groups in Latin America, it was the colonial strategies *per se* that may have allowed these groups to preserve some of their traditional forms of organisation -including the institutional ones. With the unfeasibility of the colonial power to control a vast territory, colonisers rely on the bureaucratic structures of the Amerindian groups -specially from the larger ones- to extract and control the natural and human resources. While the colonial elite were settled in more accessible places, the Amerindian groups were gradually pushed to live in remote areas. Though, the influence of the colonial power in these areas were limited on the basis of an indirect rule allowing the Amerindian groups to preserve some of their pre-colonial factors. With the conjecture that some pre-colonial factors persisted during colonisation it may be the case that -through these Amerindian rural communities- the influence of pre-colonial institutions affected the wealth of Latin American nations.

Now, having a separate life in the remote areas as a result of the indirect rule set up by the colonial power meant that the Amerindian groups had to provide a certain level of organisation in their local communities. Thus, the degree of organisation within these communities may have depended on the ability of the Amerindian groups to learnt the new political configuration in order to increase the benefits for their own groups, and the rest of the population too. However, some groups were in a better position to do so than others (Mallon 2010, p. 284). Larger Amerindian groups, such as the Incas and Aztecs, not only had the political capacity to use the new political system for their own benefit but also they had the institutional advancements to provide the basis and strategies for the development of the economic and political institutions implemented by the colonial power (e.g. Encomiendas, Mitas). For example, as the colonial power set out its settlement strategies in the Southern, Andean groups -through their local leaders- influenced Spaniards to undertake an indirect rule upon the idea that it would be more

economically viable for them (Saignes 1999, p. 66). Certainly with the expected restrictions imposed by a colonial ruler, local leaders were then given the responsibility to administer the Spanish tributes and organise the supply of labour for the mining production that was underway in this region.

When more advanced groups were influencing the strategies and expansion of the colonial power, stateless groups were constantly seeking a more resistance position towards the Europeans throughout the colonial and post-colonial periods, and most of the time in a hostile way. Located in inaccessible geographical locations, these groups -such as the Chacos groups in Paraguay and Brazil or the Aracaunos in Chile- had weaker institutional capacity by which the colonial rulers could rely on to control the large number of hunter-gathering groups scattered in these inhospitable areas. The interaction with these groups were reduced to temporary, strategic and -in most cases- failed missionary campaigns leaving little space for these small groups to influence the political system for their advantage, and thus failing to offer better benefits for their own communities.

These contrasting scenarios of the capacity of Amerindian groups surely were key in shaping the development of their local communities all along the colonial and post-colonial periods. While the communities linked to more advanced groups enjoyed periods of relative social stability, areas where stateless groups predominated were essentially characterised for being in a constant conflict. For instance, the area inhabited by the Aracauno group in Chile is known as a region of "permanent war" not just during the colonial time but also throughout the post-colonial period (Jones 1999, p.168) . As a matter of fact, stateless groups were not only constantly fighting with the Europeans but also amongst them (Schwartz and Salomon 1999, p. 460), a situation that may have created social and economic instability and, thus, underdevelopment within their communities.

The above is consistent with the recent findings on ethnic conflicts arguing that more included ethnic groups are more likely to show stability than the excluded groups (Wucherpfennig et al. 2015). Disengaged from the political context, stateless groups excluded themselves to the periphery areas of the colonial and post-colonial rules making them less peaceful than the larger Amerindian groups, and therefore, less likely to take advantage from the new political status. As a notable example from late 19th century is the Pacification War carried out in conjunction by Chile and Argentina in order to subjugate less peaceful groups located in the Pampas and Aracaunian regions.

On the other hand, all along the colonial period more advanced Amerindian groups were in a better position to even litigate with the High Courts of New Spain over more land guarantees for their local communities. The Amerindian groups located in the inner area of Central Mexico took advantage of the political system established by the colonial rule

to obtain -through legal disputes- more land titles in the hope of increasing the economic gains of their respective communities. A particular case was documented in 1550's, in the village of Huejotzingo of the Mexican state of Puebla, where native chieftains recovered "*a big expanse of land under cover of the serious dissensions that had troubled this large village*". (Francois 1963, p. 208). In addition to this, these Amerindian groups found ways to generate further income through the rent of communal property to non-Amerindians (Young 2000, p. 157). As a result, one may argue that with more economic opportunities these Amerindian groups may have redistributed these resources within their communities in the form of better public goods.

Precisely, if one may want to relate the capacity of Amerindian groups to influence and, in some ways, "support" the new political system is by looking at their role in defending their communal lands all along the colonial and post-colonial periods. In light of the possession over some of the lands through the expansion of the land-holding system of Haciendas in later colonial period, the Amerindian groups located in Central Mexico somehow managed to overcome this inflicting expansion (Schryer 2000, p. 237). For instance, in order to take hold of as many land as possible, Spanish settlers bought lands to indebted Amerindians for a small sum, or simply seized lands belonging to deceased Amerindians. A clear example was documented in mid-XVII century, in the state of Oaxaca, when a Spanish Crown's settlement inspector found irregularities in the haciendas' title deeds. Fraçois quoted the main conclusion of such Spanish inspector: "*It has been ascertained that all the land sold by notables and inhabitants of these towns, villages, and quarters did not belong to them, but to Indians dying without heirs*" (Francois 1963, p. 217). Yet, as Fraçois noted, larger groups "*showed the greatest skill in resisting big sheepmen and hacendados who wanted to force them into peonage. They obtained royal injunctions, went to court over their privileges, and cleverly played on the kindly feelings of certain of their titular protectors*" (Francois 1963, p. 219).

The political participation of Amerindian groups went even beyond their local communities. In the advent of the post-colonial period, educated Amerindian groups in Mexico were in the position to campaign with the central government in favour of preserving the institutions that sustained the economy of their local communities (e.g. *parcialidades*). Schryer (2000, p. 251) notes that during the first decades after independence "*small elite of educated natives...voiced their protest against a government policy that would abolish all native corporate rights or special institutions designed for native ethnic minorities*". These political gains may have been transferred into the well-being of their local communities.

However, the political participation of Amerindian groups was affected throughout the post-colonial era by a phenomenon that was perhaps not seen in other parts of the colonised world: *mestizaje*. Since late 18th century and onwards, all the mestizo-related

groups in Latin America achieved rapidly a dominant position amongst their societies allowing them to set up the new rules of the game. In building their countries, the mestizo groups called for a single national identity with the argument that all individuals were to be ruled under the same policies regardless their cultures, race or social status (Hill 1999). In fact, their constitutions were drafted under these principles of "equality". The political agenda led by the mestizo groups, however, implicitly disregarded the existence of a multi-ethnic structure hoping that in a later stage the Amerindian groups would integrate themselves as full "citizens" into the national projects.

Yet the so called "Indian problem" was far from being solved throughout the post-colonial period. Since some Amerindian groups played a major role in the post-independence struggles, they regained some political autonomy within their local communities. For example, after the Mexican Revolution, policies on land redistribution shaped constantly the political agenda of the central governments. A situation that throughout the 20th century gave way to the conformation of a corporatist relationship between the Amerindian communities and the central governments -a similar pattern took place in Bolivia, Ecuador, Guatemala and Peru, countries that have one of larger Amerindian groups in Latin America. In theory, this political link aimed to weaken the ethnic identity of Amerindian groups in order to integrate them into the single national identity. However, in practical terms, these corporatist policies enabled some of the Amerindian communities to have access to better social and economic benefits, and above all, to increase the space within their local communities of a certain degree of political autonomy (Deborah 1998).

By the 1980s, the rise of the "indigenous movements" marked the explicit political importance of Amerindian groups in all Latin America -specially for the larger groups. Corporativism formulated a dual identity amongst the Amerindian groups -peasants versus Amerindians- that, on one hand, in dealing with the central government, Amerindians identified themselves as peasants in order to achieve better social rights. On the other hand, within their local communities they assumed their role as Amerindians enabling them to gain local support through a network of cross communal entities (Deborah 2005, p.64). As the corporatist regimes started to get weaker, this solid local based Amerindian community network erupted by claiming -most of the time in a peaceful way- better benefits for their respective communities (Van Cott 2010). Larger Amerindian groups such as Incas in Peru or Ecuador were engaging with their central governments to demand better public goods for their communities such as bilingual education (Selverston 1995, p.131). In countries with a more dense Amerindian population, larger groups such as Aymara managed to achieve a national representation allowing them to elect a president with Amerindian origins in Bolivia in mid-2000s. These waves of Amerindian-based movements shifted the national policies from backward views towards the Amerindian

population to a debate in which the Amerindian groups are not longer seen as isolated political actors but as relevant elements to foster democracy and economic development.

4.4 Data and Methodology

To test the main hypothesis of the present study I combine two data sources. The main data source is drawn from the Latinobarometer. This is an annual public opinion survey on individuals' attitudes, values and behaviour towards democracy, economy and society, covering 18 Latin American countries. The survey has been conducted every year since mid-1990s. Although the present analysis is mainly based on the 2010's survey, I also tested the hypothesis with other annual versions after 2001, obtaining similar results. Versions prior 2001 did not gather data on ethnic identity. Such ethnicity information was required for the empirical strategy to investigate the different effects of the legacy of pre-colonial institutions according to the different type of populations (e.g. mestizo, white, indigenous, etc.).

The simple question that measures individuals' attitudes towards different contemporary political institutions in the Latinobarometer is: "*...how much trust do you have in each of the following groups/institutions...?*". The respondents are given a list of various institutions amongst them *i)* states; *ii)* national congress; *iii)* national electoral commission; and *iv)* public administration. They are then asked to rank each of these institutions in the following categorical way: 1) *a lot*; 2) *some*; 3) *a little*; and 4) *no trust*. I create four different indicators taking a value of 1 if individuals' responses are categories 1) *a lot*, 2) *some*; 3) *a little*; and zero for category 4) *no trust*. These four indicators will be my main alternative dependent variables measuring the support of individuals for contemporary political institutions.

The second data source comes from the state-level dataset prepared for the second chapter of the present thesis. I am particularly interested in using the index that measures the institutional complexity among the Amerindian population in each subnational state of Latin America -please refer to the methodological section of Chapter 2 for more details on the index. I use this subnational index to explore the main hypothesis of the present study. As such, I just have to assign to each individual of the survey its respective value of this index using the location of states as the main criteria for the merging of the data¹.

¹ Amongst the questions that the survey computes is the location of individuals disaggregated into three levels: countries, states, and cities. I select states to undertake the merging of the data. For example, explicitly speaking, all the individuals of the survey that indicated living in the state of La Paz, Bolivia, are taken the value of the index calculated for this Bolivian state. Consequently, the value on levels of pre-colonial institutions will be the same for individuals living in La Paz.

With this information at hand, I am in the position to explore the long-run effects of pre-colonial institutions on people's attitudes. First, I will assume that individuals are affected by the effects of pre-colonial institutions of the area where they live, regardless of their ethnic origins. I will call this as an overall effect of pre-colonial institutions. Second, since the legacy of pre-colonial institutions would not be the same across all the different ethnic groups, I will examine if these long-run effects vary depending on whether individuals identify as Amerindian, Mestizo, White, Blacks, others races, etc.

To explore the overall effect of pre-colonial institutions on the individuals' perception I simply use the information on pre-colonial institutions by itself on my four alternative outcomes related to the degree of support for contemporary political institutions. If the main hypothesis of the study is correct, we expect a positive association between them, indicating that areas historically inhabited by more advanced Amerindian groups -as reflected by the information on the levels of pre-colonial institutions- relate to individuals who are more likely to support the contemporary political institutions. Yet it is important to note that this effect may be biased due to the fact that individuals in more developed regions are more likely to support the political system. In consequence, the evidence throughout this Chapter may be considered cautiously.

The overall effect of pre-colonial institutions on the individuals' attitudes will be analysed by using the following logit regression model:

$$\text{Logit}(\pi_{i,s,c}) = \beta PCI_{s,c} + \zeta AmePop_{s,c} + \theta D_{i,s,c} + \lambda X_{s,c} + \alpha_c + \varepsilon_{i,s,c} \quad (4.1)$$

In equation 4.1 suscript i indicates individuals, suscript s denotes subnational states and suscript c signifies countries. π is the probability that one of our indicators measuring the support of individuals for contemporary political institutions is equal to 1. In notation form, this is the same as $\Pr(\pi = 1)$. Our subnational index of pre-colonial institutions is $PCI_{s,c}$. It is worth pointing out that in order to easily interpret the effect of our coefficient for PCI, results will be reporting marginal effects at the means. Thus, the effect would be in terms of how much $\Pr(\pi = 1)$ is predicted to change as β increases by one-unit -holding all other covariates constant. Of course, the effects of the coefficient for our set of controls would be interpreted similarly.

Turning to our equation 4.1, $AmePop_{s,c}$ controls for the share of Amerindian population in each state in Latin America. This information was drawn from the same subnational dataset where our index of pre-colonial institutions was obtained -and thus merged with the individual-level data using the same steps. This control is important since the influence of Amerindians on politics may be due to their sheer numbers instead of their pre-colonial

institutions. $\mathbf{X}_{s,c}$ is a vector of geographic and historical controls. These variables also come from such subnational dataset. As we progress in the empirical analysis I will point out the inclusion of these controls. $\mathbf{D}_{i,s,c}$ is a set of individual demographic variables on age, gender, socioeconomic status, location and educational levels of respondents. Finally, α_c is set of country fixed effects.

Yet by exploring an overall effect the analysis would then be restricted on the variation of perception of all the individuals of the survey regardless their respective ethnic identity. The assumption is that the legacy of pre-colonial institutions would be more marked on individuals who may have been more exposed to the pre-colonial norms that have persisted through the Amerindian communities as a result of their possible proximity. This may be the case for all the variants of the dominant mestizo groups (e.g. Bolivians, Mexicans, Colombians, etc.). Thus, the long-run effects of pre-colonial institutions on individuals' attitudes would vary depending on the type of population within each subnational state, and not just on an overall average effect.

Our second empirical strategy is therefore to explore the long-run effects of pre-colonial institutions on individuals' attitudes by using the information of the survey on the different ethnic identities of individuals. To do so, I proceed as follows. I implement a series of two-by-two interactions between the index of pre-colonial institutions and the different ethnic identities as classified by the Latinobarometer. Since 2001, the Latinobarometer has included the question *Which of these ethnic groups do you belong to?*, in order to find out the ethnic identity of the respondents. Yet this information has been gathered by using a generic classification. The 2010's survey, for instance, provides seven ethnic classifications: 1) asian; 2) black; 3) indigenous; 4) mestizo; 5) mulato; 6) white; and 7) other race. Ideally, information on detailed ethnic identity covering the wide range of Amerindian groups (e.g. Incas, Chibchas, Tobas, etc.) would shed better lights on this empirical examination. However, with the available information, I can recode indicators for these seven ethnic classifications². Since there is a strong reason to believe that the long-run effects of pre-colonial institutions may be different depending on the type of population, by interacting each of these indicators with the value of the index of pre-colonial institutions, I would capture the variation of people's attitudes towards the contemporary political institutions not just on the overall effect of pre-colonial factors but on the basis of the respective group of population.

To investigate the long-run effects of pre-colonial institutions on individuals' attitudes on the basis of the different ethnic identities of individuals, I use the following logit re-

²I also aggregated the smaller groups into two indicators: Asian and Other races; Mulatos and Blacks. I ended up with four indicators: Amerindian, Blacks, Mestizos, White, and Others. Yet, by using this classification, results do not improve.

gression model:

$$\begin{aligned}
\text{Logit}(\pi_{i,s,c}) = & \zeta \text{AmePop}_{s,c} + \mathbf{I}_{ASIANi,s,c} + \mathbf{I}_{BLACKi,s,c} + \mathbf{I}_{MESTIZOi,s,c} + \\
& \mathbf{I}_{MULATOi,s,c} + \mathbf{I}_{WHITi,s,c} + \mathbf{I}_{OTHERi,s,c} + \beta_1 \text{PCI}_{i,s,c} * \mathbf{I}_{AMERINDIAN,s,c} + \\
& \beta_2 \text{PCI} * \mathbf{I}_{ASIANi,s,c} + \beta_3 \text{PCI} * \mathbf{I}_{BLACKi,s,c} + \\
& \beta_4 \text{PCI} * \mathbf{I}_{MESTIZOi,s,c} + \beta_5 \text{PCI} * \mathbf{I}_{MULATOi,s,c} + \beta_6 \text{PCI} * \mathbf{I}_{WHITi,s,c} + \\
& \beta_7 \text{PCI} * \mathbf{I}_{OTHERi,s,c} + \theta \mathbf{D}_{i,s,c} + \lambda \mathbf{X}_{s,c} + \alpha_c + \varepsilon_{i,s,c}
\end{aligned} \tag{4.2}$$

In specification 4.2 our outcomes will be the same as in equation 4.1; along with the control for share of Amerindian population - $\text{AmePop}_{s,c}$ -; the two vectors of controls for individual demographic characteristics and geography plus historical factors - $\mathbf{D}_{i,s,c}$, $\mathbf{X}_{s,c}$ -; and the set of country fixed effects - α_c . The difference is that I have now included seven two-by-two interactions between the index of pre-colonial institution and the seven indicators on ethnic identity. These interactions are represented in specification 4.2 with the coefficients β_1 - β_7 . With the inclusion of these series of two-by-two interactions the long-run effects of pre-colonial institutions on people's attitudes is now affected by the changes according to the ethnic identity of individuals. The most significant changes can be summarised as follows:

- i) There is no longer an overall effect of PCI since now it depends upon the different ethnic identities as measured by our seven indicators.
- ii) The effect of PCI on the basis of the ethnic identities will be obtained, as above, by computing the marginal effects at the means. This will allow us to see how much $\text{Pr}(\pi = 1)$ changes as PCI increases by one-unit depending on our seven indicators on ethnicities - holding all other covariates constant.

4.4.1 Descriptive statistics

Panel A in Table 4.1 shows summary statistics on our four alternative outcomes measuring the degree of support for contemporary political institutions for the whole sample³ These indicators report the individuals within each state of the 17 Latin American countries

³It is worth noting that the 19, 204 observations come from 253 states belonging to 17 Latin American countries. The number of states differs from the subnational dataset prepared for the first chapter - standing at 324 states- due to data collection criteria implemented by the Latinobarometer. For example, for some countries the criteria was to use major regions rather than the actual states. That was the case for Brazil and Colombia. For this, I simply average the values from the subnational dataset depending on the respective region applied in the survey. The aim was to synchronise the variables from both data sources.

supporting the *i)* states; *ii)* national congress; *iii)* national electoral commission; and *iv)* public administration. The percentage of individuals supporting these four alternative national institutions range between 73 to 79, suggesting that almost three in one individuals in Latin America seem to support this selected set of contemporary political institutions. Note that the strong degree of support for political institutions may be overestimated due to -but not only- country-specific national policies influencing upwards people's attitudes -an issue that we want to tackle with the use of country-fixed effects.

The summary statistics of our main variable of interest is also reported in Table 4.1. After assigning to each respondent of the survey the respective value of the index of pre-colonial institutions from his/her subnational state, the variable on pre-colonial institutions gets a mean value of .62. This relative small number is not surprising since most of the values of the index are towards zero. Yet this number hides an important variation. For example, for the individuals in Ecuador or Bolivia the total mean of their values of pre-colonial institutions is around 1 and 3, respectively. This variable by itself, thus, will be capturing the overall long-run effects of pre-colonial institutions on the four alternative outcomes on the degree of contemporary political institutions.

Demographic and socioeconomic controls are also reported in Table 4.1. As a noteworthy point on this information is that around of 95% of the individuals come from urban areas. This is of relevance for the empirical investigation since, as outlined before, mestizos tend to live in larger cities where the influence of pre-colonial norms may not be as marked as in rural areas. Yet I may argue that if the analysis favours the main hypothesis, I would then assume that by increasing the percentage of individuals from rural areas the results would be more robust -specially when I analyse the effect according to the different ethnic composition. The rest of the rows report summary statistics on the following controls: share of Amerindian population, demography and historical factors. All these last set of controls come from the subnational dataset prepared in Chapter 2 as well. Thus, the information of these controls was assigned to each respondents in the same way as with the values of the index of pre-colonial institutions. Note that due to the missing data on some of our historical controls, the sample size slightly reduces, yet the remaining data still allows us to exploit nearly 90% of the total number of observations in the analysis.

Panel B in Table 4.1 presents summary statistics of the indicators on the various ethnic identities as classified by the 2010's Latinobarometer. The means of the indicators show the fraction of the population in the survey data of each ethnic identity. Nearly half of the population identified itself as having mestizo origins -in other words with a mix of European and Amerindian origins-. The second largest ethnic group is the white population, with nearly 30% of individuals having European origins. The remaining individuals indicated having either Asian (0.08%), Indigenous (8%), Black(3%), Mulato (5%) or Other (7%)

Table 4.1: Summary statistics

Variables	Observations	Mean	Std. Dev.	Min.	Max.
Panel A					
<i>Main outcomes:</i>					
Indicator supporting the State	18461	0.7928606	0.4052673	0	1
Indicator supporting the Congress	18446	0.7343598	0.4416855	0	1
Indicator supporting the Electoral Body	18231	0.7916187	0.406162	0	1
Indicator supporting the Pub. Admin.	18322	0.7722956	0.4193623	0	1
<i>Main Independent variable:</i>					
Pre-colonial Institutions (PCI)	19204	0.627872	1.118814	0	3.997829
<i>Individual-level controls:</i>					
Socioeconomic level	19204	2.695428	0.89061	1	5
Rural	19204	0.046084	0.209673	0	1
Urban	19204	0.953916	0.209673	0	1
Illiterate	19204	0.098834	0.298446	0	1
Incomplete primary	19204	0.194595	0.395899	0	1
Complete primary	19204	0.174859	0.379857	0	1
Incomplete secondary	19204	0.136899	0.343749	0	1
Complete secondary	19204	0.225786	0.41811	0	1
Incomplete high	19204	0.082639	0.275343	0	1
Complete high	19204	0.086388	0.280944	0	1
Male	19204	0.480733	0.499642	0	1
Female	19204	0.519267	0.499642	0	1
Age	19204	40.18517	16.5252	16	96
Age squared	19204	1887.916	1517.718	256	9216
<i>Subnational level controls:</i>					
Share of Amerindian population	19204	0.088615	0.169105	0.000163	0.941406
Latitude	19204	16.86296	10.08407	0.015341	52.47901
Altitude	19204	0.736463	0.950766	0	4.049
Temperature	19160	20.26219	5.131414	5.028432	27.43242
Land area (km. sq.)	19204	54777.23	92461.33	63.1	605709.2
Distance to capital	19204	316.1279	351.2065	0	2103.238
Inverse distance to the sea	19160	0.899102	0.093606	0.618014	0.99904
Landlocked	19204	0.515101	0.499785	0	1
Oil and gas resources	19204	0.168663	0.374464	0	1
Gold and silver resources	19204	0.235993	0.424629	0	1
Population density (logs)	19204	4.198514	1.721223	-1.96108	10.98031
Pre-colonial population density	17827	0.763271	2.562362	-9.5814	5.972121
Other colonial economic activities	16677	0.78977	0.407484	0	1
Mining during colonial period	16677	0.191761	0.393698	0	1
Plantations during colonial period	16677	0.121785	0.327047	0	1
Panel B					
<i>Indicators:</i>					
Asian	19204	0.008071	0.089479	0	1
Black	19204	0.034055	0.181376	0	1
Amerindiand	19204	0.083264	0.276288	0	1
Mestizo	19204	0.439856	0.496383	0	1
Mulato	19204	0.052489	0.223017	0	1
White	19204	0.299052	0.457855	0	1
Other	19204	0.073995	0.26177	0	1

ethnic origins. Quite remarkably is that the mean of our indicator of interest measuring the fraction of the Amerindian population (Indigenous) seems to have a similar value as the mean of the control variable on the share of Amerindian population drawn from statistics quantifying the total population within each country.

4.5 Empirical Results

The first empirical strategy explores the overall long-run effects of pre-colonial institutions on the degree of support for contemporary political institutions. The assumption is that individuals are affected by the effects of pre-colonial institutions of the region where they live regardless of their ethnic origins. To start this examination, in Table 4.2, I use as our dependent variable the indicator on whether individuals support the state or not. Since our outcome takes a binary form, I report Logit estimators all along the specifications. I run four separate regressions by including all through our variable of interest on pre-colonial institutions. Controls will be added gradually so that our main coefficient is under an in-depth examination. As mentioned before, to easily interpret the results, I am reporting marginal effects at the means.

In model (1) from Table 4.2, the variable on pre-colonial institutions is included by itself, along with country-fixed effects. I then cluster standard errors at country-level all through the specifications⁴. As mentioned before, the inclusion of fixed effects is important due to the fact that key country-specific unobservable factors may be affecting our results. For instance, some -but not only- of the effects that I am interested in controlling for are the policies implemented by the central governments that may have had a strong influence on people's attitudes. The assumption is that through these national policies the presence of central governments reaches out remote communities affecting the perception of individuals towards the role of national political institutions. If this may be the case, individuals would support the national political institutions regardless what kind of legacy of pre-colonial institutions they have. So, by not controlling for these hard-to-account-for- unobservables factors, the effect of our variable of interest may be potentially biased. As it stands, the marginal effects for our coefficient of interest $-\beta$ - show a positive and significant probability of success conditional on key country-specific unobservable factors. As PCI increases by one-unit the change in the probability that individuals would support the state is about 2.5%. This result supports already the main hypothesis of the present study.

In the second regression from Table 4.2, the control for the total share of Amerindian population in each state is added. By adding this control the coefficient on pre-colonial

⁴I also cluster standard errors at both subnational state and country levels. Results are not affected.

Table 4.2: Pre-Colonial Institutions and Population Supporting the State

Dependent variable: Population Supporting the State				
Variables:	(1)	(2)	(3)	(4)
Pre-Colonial Institutions	0.0251* (0.0144)	0.0248* (0.0136)	0.0246* (0.0147)	0.0331*** (0.0103)
Share of Amerindian pop.		0.00447 (0.0617)	0.00979 (0.0611)	0.0450 (0.0699)
Log population density			-0.00362 (0.00346)	0.00321 (0.00375)
Demographic controls:				
Socio economic level (ref. Very good)				
Good			-0.0130 (0.00959)	-0.0146 (0.0108)
Not bad			-0.0344*** (0.0118)	-0.0470*** (0.0111)
Bad			-0.0525** (0.0218)	-0.0713*** (0.0216)
Very bad			-0.0732** (0.0335)	-0.0916*** (0.0286)
Area (ref. Rural)			-0.0271 (0.0199)	-0.0375** (0.0179)
Education (ref. Illiterate)				
Incomplete primary			0.00695 (0.0153)	3.24e-05 (0.0180)
Complete primary			-0.00294 (0.0219)	-0.00395 (0.0258)
Incomplete Secondary, technical			0.0125 (0.0191)	0.00653 (0.0216)
Complete Secondary, technical			0.0153 (0.0201)	0.00751 (0.0232)
Incomplete high			0.0302 (0.0237)	0.0190 (0.0260)
Complete high			0.0321 (0.0231)	0.0212 (0.0261)
Gender (ref. Male)			-0.00327 (0.00585)	-0.00278 (0.00644)
Age			-0.00272** (0.00114)	-0.00197* (0.00109)
Age-squared			2.62e-05** (1.16e-05)	1.85e-05* (1.11e-05)
Geography, location, and natural resources:				
Latitude				-0.000933 (0.00197)
Altitude (km.)				-0.00824 (0.00671)
Temperature (Celsius)				0.00262 (0.00224)
Land area (sq. km.)				3.70e-08 (9.27e-08)
Distance to capital (km.)				6.18e-05* (3.19e-05)
Inverse distance to coast				0.152 (0.132)
Landlocked dummy				8.87e-05 (0.0182)
Oil & Gas dummy				0.0248 (0.0288)
Gold & Silver dummy				0.0381** (0.0157)
Historical controls: colonial & pre-colonial				
Pre-colonial population density				-0.00149 (0.00400)
Other colonial activities				-0.0119 (0.0294)
Mining				-0.00179 (0.0307)
Plantations				-0.0211 (0.0422)
Country Fixed Effects	YES	YES	YES	YES
Observations	18,461	18,461	18,461	15,864

Notes: Cluster standard errors at country levels are in brackets
*** p<0.01, ** p<0.05, * p<0.1

institutions is not affected. This is not surprising since the coefficient on the share of Amerindian population is insignificant. It seems that by operationalising this control at the individual-level the expected effect is unclear. Regression (3) from Table 4.2 includes our control for population density, along with a set of individual characteristics on age, age squared, gender, education, areas and socioeconomic level. As it can be seen, the marginal effects for the coefficient β is stable. In model (4), I control for geography and historical factors. As it stands, our coefficient of interest on pre-colonial institutions remains positive and very significant. Though, the effect is relatively small. The coefficient takes a value of 3.3%. If we consider the average value of the indicator measuring the degree of support for the state of 79% with a standard deviation of 40%, then an increase of one-unit of our index of pre-colonial institutions would only lead to 0.08 standard deviation change in the probability of individuals supporting the state.

Table 4.3 reports our results on the effects between the legacy of pre-colonial institutions and the four alternative indicators on contemporary political institutions. The regressions use specification (4) from Table 4.2 as our baseline model, and the outcomes on whether individuals support the state, national congress, electoral bodies and public administration. The results are consistent with the main hypothesis of the present chapter. In all the regressions the marginal effects for our coefficients on pre-colonial institutions show the correct sign and very high significant levels. Interestingly, the marginal effects for the coefficients β in models (1)-(3) reflect a higher success of individuals supporting the state, congress and electoral bodies, as compared to the coefficient for the specification on supporting public administration. This may suggest that areas historically inhabited by institutionally advanced Amerindian groups tend to have a stronger perception on the role of the main central authorities over any other type of governmental entity. It is also worth remarking the results in model (3). The results of this model may go in line with the recent findings emphasizing the growing role of the Amerindian population in shaping the electoral systems in Latin America (Van Cott 2010). It may suggest that the degree of political salient of ethnic cleavages may be rooted in the legacy of institutional complexity of Amerindian groups in Latin America, holding constant the effect of the share of Amerindian population.

Yet what we have been capturing so far is an overall long-run effect of pre-colonial institutions. This implies that all the population as a whole -Amerindians or non-Amerindians- may have been affected by an effect that is meant to have persisted -principally- in the remote Amerindian communities. Such a conclusion may be, however, put into question easily, considering that the ethnic composition in Latin America since the 18th century came to be significantly transformed by particularly what is today the different variant of

Table 4.3: Pre-Colonial Institutions and Contemporary Political Institutions

Dependent variables:	Supporting State (1)	Supporting Congress (2)	Supporting Electoral Body (3)	Supporting Admin. Public (4)
Pre-Colonial Institutions	0.0331*** (0.0103)	0.0375*** (0.0122)	0.0269*** (0.00726)	0.0163*** (0.00455)
<i>Controls for:</i>				
Share of Amerindian pop.	YES	YES	YES	YES
Log population density	YES	YES	YES	YES
Demographic controls	YES	YES	YES	YES
Geography, location, and natural resources	YES	YES	YES	YES
Historical controls: colonial & pre-colonial	YES	YES	YES	YES
Country Fixed Effects	YES	YES	YES	YES
Observations	15,864	15,817	15,638	15,741

Notes: Cluster standard errors at country levels are in brackets

*** p<0.01, ** p<0.05, * p<0.1

the dominant groups called mestizos. These dominant groups have mostly been shaped by the post-colonial political institutional features making less likely that their present-day political attitudes may have been influenced by the long-run effects of pre-colonial institutions. In addition to this large group, there are other ethnic compositions that are of relevance for some Latin American countries. For instance, the population that came through the slave trade is nowadays an important ethnic composition for countries like Colombia, Venezuela or Brazil. Moreover, the Europeans that settled permanently, and that today are meant to be related as being part of the white group, is more predominant in some countries in South America such as Chile and Argentina.

Thus, I next focus on the long-run effects of pre-colonial institutions on the basis of the different ethnic compositions in Latin America. Our main motivation to carry out such approach is based on the possibility that the legacy of pre-colonial factors may depend on what type of population we are dealing with. To conduct such examination I will use the ethnic identities as classified by our survey data in the way that was explained in the previous section on data and methodology.

Table 4.4 reports the main results regarding the interactions of the different ethnic identities and our measure of long-run effects of pre-colonial institutions. We are using our four different alternative outcomes on the degree of support for contemporary political institutions, along with the battery of controls as of model (4) from Table 4.2. The only difference is that six two-by-two interactions $-\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7$, plus their respective indicators $-\mathbf{I}_{ASIANi,s,c}, \mathbf{I}_{BLACKi,s,c}, \mathbf{I}_{AMERINDIAN,s,c}, \mathbf{I}_{MESTIZOI,s,c}, \mathbf{I}_{MULATOi,s,c}, \mathbf{I}_{WHITI,s,c}, \mathbf{I}_{ASIANi,s,c}, \mathbf{I}_{OTHERi,s,c}$ are included on these specifications.

The results are meaningful. The long-run effects of pre-colonial institutions do seem

Table 4.4: Population groups

Dependent variables:	Supporting State (1)	Supporting Congress (2)	Supporting Electoral Body (3)	Supporting Admin. Public (4)
Marginal effects of PCI*dummy for Amerindians	0.0330*** (0.0094)	0.0359 (0.0243)	0.0256** (0.0109)	0.0093 (.0069)
Marginal effects of PCI*dummy for Asians	0.0565 (0.0392)	0.0912** (0.0398)	0.2169*** (0.0818)	0.0852** (.0459)
Marginal effects of PCI*dummy for Blacks	0.0148 (0.0269)	0.0323 (0.0268)	-0.0067 (0.0232)	0.0167 (.0205)
Marginal effects of PCI*dummyfor Mestizos	0.0352** (0.0180)	0.0345*** (0.009)	0.0256*** (0.0073)	0.0147*** (.0039)
Marginal effects of PCI*dummy for Mulatos	0.0225 (0.0194)	0.0359 (0.0239)	0.0104 (0.0183)	0.0401** (0.0175)
Marginal effects of PCI*dummy for Whites	0.0377*** (0.0117)	0.0333*** (0.0128)	0.0365*** (0.0142)	0.0175* (0.0093)
Marginal effects of PCI*dummy for Other Races	0.0251* (0.0129)	0.0491*** (0.015)	0.0366*** (0.0099)	0.0237 (0.0156)
p-value Wald test	0.000	0.5199	0.0114	0.0025
<i>Controls for:</i>				
Share of Amerindian pop.	YES	YES	YES	YES
Log population density	YES	YES	YES	YES
Demographic controls	YES	YES	YES	YES
Geography, location, and natural resources	YES	YES	YES	YES
Historical controls: colonial & pre-colonial	YES	YES	YES	YES
Country Fixed Effects	YES	YES	YES	YES
Observations	15,714	15,665	15,491	15,590

Notes: Cluster standard errors at country levels are in brackets
*** p<0.01, ** p<0.05, * p<0.1

to vary depending on the major ethnic identity of individuals yet this difference is not statistically significant in model 2 as reported by the p-value for the Wald test. The marginal effect of PCI through the Amerindian population is only positive and significant in model (1) and (3). The effects through the Mestizo and White groups are always positive and significant. Interestingly, the marginal effects of Amerindians, Mestizos, and Whites are about the same. Perhaps, the ethnic identity with the lesser connection with the pre-colonial institutions is the Black and Mulato groups. This may suggest that Mulato and Blacks groups tend to be more excluded from the political sphere, and thus the effects of PCI are the least clear. Though, this result may be due to the fact that these groups have the lowest proportion of population.

Overall, the results of Table 4.4 suggests that the effects of PCI through the Amerindian population is not as big or different than the rest of the population, particularly with respect to the largest groups such as Mestizos and Whites. Yet we may say that the effects of PCI tend to be different depending on the type of population. This may provide evidence of some positive externalities of the pre-colonial institutions coming from the Amerindian communities towards other ethnic groups in Latin America.

4.5.1 Robustness checks

The robustness of the results are tested against various alternative approaches. First, I construct an indicator taking a value of one if the four indicators on supporting the

Table 4.5: Robustness checks (Logit vs Probit)

Dependent variable: Indicator measuring the degree of support for all the contemporary institutions				
	Logit (1)	Logit (2)	Probit (3)	Probit (4)
Pre-Colonial Institutions	0.0345** (0.0143)		0.0347** (0.0142)	
Marginal effects of PCI*dummy for Amerindians		0.0255 (0.0176)		0.0252 (0.0175)
Marginal effects of PCI*dummy for Asians		0.1391** (0.0629)		0.1368** (0.0559)
Marginal effects of PCI*dummy for Blacks		0.0365 (0.0324)		0.0373 (0.0327)
Marginal effects of PCI*dummy for Mestizos		0.0364** (0.0159)		0.0366** (0.0159)
Marginal effects of PCI*dummy for Mulatos		0.0341 (0.0213)		0.0333 (0.0213)
Marginal effects of PCI*dummy for Whites		0.0299* (0.0159)		0.0304* (0.0159)
Marginal effects of PCI*dummy for Other Races		0.0355** (0.0139)		0.0358*** (0.0138)
p-value Wald test		0.2409		0.1686
<i>Controls for:</i>				
Share of Amerindian pop.	YES	YES	YES	YES
Log population density	YES	YES	YES	YES
Demographic controls	YES	YES	YES	YES
Geography, location, and natural resources	YES	YES	YES	YES
Historical controls: colonial & pre-colonial	YES	YES	YES	YES
Country Fixed Effects	YES	YES	YES	YES
Observations	16,451	16,294	16,451	16,294
Notes: Cluster standard errors at country levels are in brackets				
*** p<0.01, ** p<0.05, * p<0.1				

state, congress, electoral body, and public administration are all equal to one, and zero otherwise. In models (1)-(2) of Table 4.5, we report Logit estimators. These specifications are using the same controls as of specification (4) from Table 4.2. Though, in model (1) the variable on pre-colonial institutions is included by itself so its coefficient will be capturing the overall effect as specified in our first empirical strategy. As the results stand, by aggregating our outcome to report individuals supporting all the four contemporary political institutions, Logit estimators still confirm our results from Tables 4.2-4.4.

In model (2), I report results using the second empirical strategy by including the seven two-by-two interactions plus the indicators for each ethnic identity. The results once again show a marked variation of the effect of PCI depending on the major ethnic identities in Latin America, though the difference is not statistically significant as reported by the p-value for the Wald test.

Second, while previous results were obtained through Logit estimators, in models (3)-(4) of Table 4.5 we report results using Probit estimators. While model (3) shows the overall effect, in model (4) the seven two-by-two interactions are added. The results are essentially held.

4.6 Conclusions

The present study investigated the mechanisms of the long-run effects of pre-colonial institutional advancements. According to our historical context, the colonial and post-colonial strategies implemented on the basis of an indirect rule plus the political advancements of Amerindian groups have worked in tandem allowing the larger Amerindian groups to support the new political systems in order to gain more political and economic benefits for their respective groups. This mechanism has allowed, on one hand, the effects of pre-colonial institutions to persist throughout the colonial and post-colonial periods, and on the other, to affect socioeconomic development outcomes in the long-run.

To shed lights on this mechanism, I combined novel data measuring the levels of pre-colonial institutions in each subnational state in Latin America and individual-level data on people's attitudes. The main empirical results showed that areas with a history of more advanced pre-colonial institutions relate to a higher probability of population supporting contemporary political institutions. Moreover, my results also showed that the effect of PCI may vary depending on the major ethnic identities given some evidence of some positive externalities coming from the Amerindian communities. Overall, this suggests that institutionally-advanced pre-colonial ethnic groups have been in a better position to process through a certain degree of support the new political systems in ways that would benefit their respective ethnic groups.

Chapter 5

Conclusions

5.1 Summary of key findings of the thesis

The present doctoral thesis studied the relationship between pre-colonial institutions and long-run development in the context of Latin American countries. The evidence provided throughout the thesis contribute to the existing literature in place for the case of the African nations: Latin American pre-colonial institutions are relevant predictors of present-day measures of socioeconomic development.

What remained after colonisation 500 years ago was not a mirror image of European society but a new reality where pre-colonial culture and institutions survived. The historical narrative underlined that pre-colonial institutions survived to our days thanks to the existence of largely self-governed Amerindian communities in rural Latin America. The existence of these Amerindian communities, inadvertently, kept alive the notion of large pre-colonial nations in the places where these existed. The central hypothesis of the present doctoral thesis is that these institutional factors continue to exert an influence on socioeconomic development to our days.

The main empirical findings of the present thesis support the above hypothesis. Using a cross-section of 324 subnational states in Latin America, Chapter 2 uncovered a strong positive effect between pre-colonial institutions and various socioeconomic outcomes –an effect that has been mainly transmitted through the Amerindian population. Results were tested against various alternatives such as controlling for country-fixed effects which allowed me to remove national factors than otherwise may have been potentially affected my estimates. In the spirit to dispel other major forces driving the variation of development over the long-run, I also controlled for other important factors related to geographical fundamentals, as well as historical events such as the type of economic activity in place in the colonial period and the economic and social profile of pre-colonial societies –in addition to the institutional complexity of ethnic groups. Lastly, I also showed that the

influence of pre-colonial institutions in rural areas is far stronger providing evidence about the historical account I gave for the transmission of pre-colonial factors.

Chapter 3 took forward the study of the association between pre-colonial institutions and present-day economic development in Latin America. I used historical ethnic homelands as my main unit of analysis. This involved relying on GIS techniques in order to digitise historiographical maps which allowed me to identify the geospatial location of ethnic homelands in Latin America as of the XVI century. This approach essentially confirmed my previous findings. I showed that ethnic homelands inhabited by more advanced ethnic groups -as measured by their levels of institutional complexity- relate to better economic development today -as reflected by satellite light density data.

Finally, in Chapter 4, I studied a mechanism linking the persistency of pre-colonial institutions over the long-run regarding the ethnic political capacity and the main strategy of dominance implemented by the colonial and post-colonial systems. As noted in Chapters 2 and 3, the political capacity of ethnic groups tended to be very heterogeneous at the time of contact with the Europeans. As such, after the conquest, larger ethnic groups were more likely to have endured the new systems introduced via the colonial and post-colonial regimes. Moreover, and in front of key factor endowments, the type of colonial system based on indirect strategies of dominance gave rise to numerous self-governed Amerindian communities all along Latin America. These historical forces may have given way in areas traditionally inhabited by larger groups to some degree of "support" to the new political systems resulting in better social and economic benefits for both the Amerindian communities and the population at large. The Chapter undertook an empirical strategy combining our index of pre-colonial institutions with individual-level survey data on people's attitudes. Results showed that areas with a history of more advanced pre-colonial institutions increase the probability of individuals supporting present-day political institutions. This degree of "support" was then taken as a possible explanation regarding this mechanism linking the persistency of pre-colonial institutions over the long-run.

5.2 Challenges and potential avenues for future research

The present doctoral thesis has not been possible without the overwhelming contributions in place in the literature as well as the colossal work of other areas of knowledge outside economics. Anthropology, history, and political science helped putting in the right place the main hypothesis of the thesis.

While enormous efforts were made in order to dissipate possible biases, some key empirical challenges remained to be addressed, for which case one may want to consider them as potential avenues for future research.

As one of the most relevant challenge is regarding the possible externalities expelled by the pre-colonial factors towards the non-Amerindian population. As you may recall, the effect that I introduced all along the present thesis has mainly been carried through the Amerindian population – as pointed out in Chapter 2. Thus, considering that all the dependent variables at my disposal have measured the socioeconomic levels for the whole population in each subnational state, and not for its Amerindian component, one may want to question whether some pre-colonial externalities have taken place to the rest of the population, as well as the extent to which this has been done. One alternative way to start making progress on this is to use information on socioeconomic outcomes for the different ethnic composition in Latin America (e.g. whites, blacks, mestizos, indigenous, etc.). By estimating econometrically the association between these separate outcomes corresponding to one of the different ethnic composition, and our measure of pre-colonial institutions, perhaps, some systematic heterogeneity in a single or more ethnic groups may be revealed.

I would also emphasise the relevance of digging more into the mechanisms linking the persistency of pre-colonial institutions. While the present thesis offered a possible explanation on this, the effects are less clear. Perhaps, one of the weaknesses of this investigation -undertaken in Chapter 4- is the way that our main variable measuring the pre-colonial institutions was operationalised, taken it from a subnational level to an individual setting. One possible alternative is to direct the attention to a mechanism that researchers may find less problematic to validate both the historical account and the empirical approach. A potential mechanism is that related to the institutional capacity of Amerindian groups to organize themselves and defend their interests in front of colonial and post-colonial regimes.

Bibliography

Acemoglu, D., I. N. Chaves, P. Osafo-Kwaako, and J. A. Robinson (2013). Indirect Rule and State Weakness in Africa: Sierra Leone in Comparative Perspective. *African Successes: Sustainable Growth* 4, 1–32.

Acemoglu, D. and M. Dell (2010). Productivity Differences Between and Within Countries. *American Economic Journal: Macroeconomics* 2(1), 169–188.

Acemoglu, D., S. Johnson, and J. A. Robinson (2005). Institutions as the Fundamental Cause of Long-Run Growth,. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*, pp. pp. 385–472. Elsevier.

Acemoglu, D., J. Simon, and J. A. Robinson (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *The American Economic Review* 91(5), 1369–1401.

Acemoglu, D., J. Simon, and J. A. Robinson (2002). Reversal of Fortune : Geography and Institutions in the Making of the Modern World Income. *The Quarterly Journal of Economics* 117(4), 1231–1294.

Alsan, M. (2015). the Effect of the Tsetse Fly on African Development. *American Economic Review* 105(1), 382–410.

Angeles, L. (2011). Institutions, Property Rights, and Economic Development in Historical Perspective. *Kyklos* 64(2), 157–177.

Banerjee, A. V. and L. Iyer (2005). History, Institutions and Economic Performance: The Legacy of Land Tenure Systems in India. *The American Economic Review* 95(4), 1190–1213.

Bockstette, V., A. Chanda, L. Putterman, S. Journal, and N. Dec (2002). States and Markets : The Advantage of an Early Start. *Journal of Economic Growth* 7(4), 347–369.

Bruhn, M. and F. A. Gallego (2012). Good, Bad, and Ugly Colonial Activities: Do They Matter for Economic Development? *Review of Economics and Statistics* 94(May), 433–461.

- Burkholder, M. A. and L. L. Johnson (1998). *Colonial Latin America* (Third Edit ed.). Oxford: Oxford University Press.
- Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011). Robust Inference With Multiway Clustering. *Journal of Business & Economic Statistics* 29(2), 238–249.
- Chanda, A., C. Justin Cook, and L. Putterman (2014). Persistence of fortune: Accounting for population movements, there was no Post-Columbian reversal. *American Economic Journal: Macroeconomics* 6(3), 1–28.
- Comin, D., W. Easterly, and E. Gong (2010). Was the Wealth of Nations Determined in 1000 BC? *American Economic Journal: Macroeconomics* 2(3), 65–97.
- Deborah, J. Y. (1998). Contesting Citizenship : Indigenous Movements and Democracy in Latin America. *Comparative Politics* 31(1), 23–42.
- Deborah, J. Y. (2005). *Contesting Citizenship in Latin America: The Rise of Indigenous Movements and the Postliberal Challenge*. New York: Cambridge University Press.
- Dell, M. (2010). The Persistent Effects of Peru’s Mining Mita. *Econometrica* 78(6), 1863–1903.
- Diamond, J. (1997). *Guns, germs, and steel*. London: W. W. Norton & Company.
- Engerman, S. L. and K. L. Sokoloff (1997). Factor endowments, institutions, and differential paths of growth among new world economics: a view from economic historians of the United States. In S. Harber (Ed.), *How Latin America Fell Behind*, pp. 260–304. Stanford: Stanford University Press.
- Fenske, J. (2014). Ecology, trade, and states in pre-colonial africa. *Journal of the European Economic Association* 12(3), 612–640.
- Francois, C. (1963). *Land and Society in Colonial Mexico: The Great Hacienda*. Berkeley: University of California Press.
- Gennaioli, N., R. La Porta, F. Lopez De Silanes, and A. Shleifer (2014). Growth in regions. *Journal of Economic Growth* 19(3), 259–309.
- Gennaioli, N. and I. Rainer (2007). The modern impact of precolonial centralization in Africa. *Journal of Economic Growth* 12(3), 185–234.
- Hall, R. E. and C. I. Jones (1999). Why Do Some Countries Produce So Much More. *The Quarterly Journal of Economics* 114(1), 83–116.

- Henderson, J. V., A. Storeygard, and D. N. Weil (2012). Measuring Economic Growth From Outer Space. *The American economic review* 102(2), 994–1028.
- Hill, J. (1999). Indigenous peoples and the rise of independent nation-states in lowland South America. In F. Salomon and S. B. Schwartz (Eds.), *The Cambridge History of the Native Peoples of the Americas*. New York: Cambridge University Press.
- Huillery, E. (2011). The impact of European settlement within French West Africa: Did precolonial prosperous areas fall behind? *Journal of African Economies* 20(2), 263–311.
- Jones, K. L. (1999). Warfare, reorganization, and readaptation at the margins of Spanish rule: The Southern margin (1573–1882). In F. Salomon and S. B. Schwartz (Eds.), *The Cambridge History of the Native Peoples of the Americas*. New York: Cambridge University Press.
- Kiszewski, A., A. Mellinger, A. Spielman, P. Malaney, S. E. Sachs, and J. Sachs (2004). A global index representing the stability of malaria transmission. *American Journal of Tropical Medicine and Hygiene* 70(5), 486–498.
- Lang, J. (1975). *Conquest and Commerce: Spain and England in the Americas*. New York: Academic Press.
- Mallon, F. E. (2010). Indigenous peoples and nation-states in Spanish America, 1780–2000. In J. C. Moya (Ed.), *The Oxford Handbook of Latin America History*. New York: Oxford University Press.
- Maloney, W. F. and F. Valencia (2015). The Persistence of (Subnational) Fortune. *The Economic Journal*.
- Mann, C. C. (2005). *1491: The Americas Before Columbus*. London: Granta Publications.
- Michalopoulos, S. and E. Papaioannou (2013). Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica* 81(1), 113–152.
- Michalopoulos, S. and E. Papaioannou (2014). National Institutions and Subnational Development in Africa. *The Quarterly Journal of Economics* 129(1), 151–213.
- Murdock, G. P. (1951). *Outline of South American Culture*. New Haven: Human Relations Areas Files Press.
- Murdock, G. P. (1967). Ethnographic Atlas: A Summary. *Ethnology* 6(2), 109–236.
- Murdock, G. P. and T. J. O’Learly (1975). *Ethnographic Bibliography of North America* (4th Editio ed.). New Haven: Human Relations Areas Files Press.

- North, D. C. (1981). *Structure and change in economic history*. New York: W. W. Norton & Company.
- North, D. C. (1990). *Institutions, institutional change, and economic performance*. Cambridge: Cambridge University Press.
- Numm, N. (2008). The long-term effects of Africa's Slave Trades. *The Quarterly Journal of Economics* 123(1).
- Numm, N. and L. Wantchekon (2011). The slave trade and the origins of Mistrust in Africa. *American Economic Review* 101(7), 3221–3252.
- Oto-Peralías, D. and D. Romero-Ávila (2016). The economic consequences of the Spanish Reconquest : the long-term effects of Medieval conquest and colonization. *Journal of Economic Growth*, 1–56.
- Psacharopoulos, G. and H. A. Patrinos (1994). *Indigenous People and Poverty in Latin America*. Washington: The World Bank.
- Ramankutty, N., J. A. Foley, J. Norman, and K. McSweeney (2002). The global distribution of cultivable lands: Current patterns and sensitivity to possible climate change. *Global Ecology and Biogeography* 11(5), 377–392.
- Saignes, T. (1999). The colonial condition in the Quechua-Aymara heartland (1570-1780). In F. Salomon and S. B. Schwartz (Eds.), *The Cambridge History of the Native Peoples of the Americas*. New York: Cambridge University Press.
- Schryer, F. (2000). Native peoples of Central Mexico since independence. In R. Adams and M. MacLeod (Eds.), *The Cambridge History of the Native Peoples of the Americas*. Cambridge: Cambridge University Press.
- Schwartz, S. B. and F. Salomon (1999). New peoples and new kinds of people: adaptation, readjustment, and ethnogenesis in South America indigenous societies. In F. Salomon and S. B. Schwartz (Eds.), *The Cambridge History of the Native Peoples of the Americas*. New York: Cambridge University Press.
- Silverston, M. H. (1995). The politics of culture: indigenous peoples and the state in Ecuador. In V. C. D. Lee (Ed.), *Indigenous Peoples and Democracy in Latin America*. New York: St. Martin's Press.
- Service, E. R. (1971). *Primitive social organization: an evolutionary perspective* (2nd editio ed.). New York: Random House.

- Sokoloff, K. L. and S. L. Engerman (2011). History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World. *Journal of Economic Perspectives* 14(3), 217–232.
- Van Cott, D. L. (2010). Indigenous Peoples’ Politics in Latin America. *Annual Review of Political Science* 13(1), 385–405.
- Wucherpfennig, J., P. M. Hunziker, and L.-e. Cederman (2015). Who Inherits the State ? Colonial Rule and Post-Colonial Conflict. *American Journal of Political Science*, 1–17.
- Young, E. V. (2000). The indigenous people fo Western Mexico from the Spanish invasion to the present. In R. Adams and J. M. MacLeod (Eds.), *The Cambridge History of the Native Peoples of the Americas*. Cambridge: Cambridge University Press.

Appendix A

Appendix –*Chapter 2*–

Table A.1: Ethnic groups

Panel A: Ethnic groups with similar or same names in the Ethnographic Atlas and in the Census		
Alacalufe, Apinayé, Aymara, Bacairi, Bororo, Botocudo, Bribri, Cakchique, Camayura, Caraja, Chacobo, Chamacoco, Chichimec, Chinantec, Chiriguano, Chorote, Chorti, Cocama, Cubeo, Curipaco, Guató, Huichol, Kuikuru, Lenca, Macusi, Mam, Mapuche, Mataco, Maya, Mazateco, Mixe, Mixteco, Mundurucu, Nambicuar, Ona, Palikur, Papago, Paressi, Piaroa, Puna, Popolucá, Q'iche, Sanema, Shavante, Sherente, Seri, Siriono, Tarahumar, Tipirapé, Tehuelche, Tenetehar, Toba, Totonac, Trumai, Tupinamba, Tzeltal, Umutina, Waikwa, Wuitoto, Yagua, Yanomamo, Yuki, Zapotec, Zoque		
Panel B: Ethnic groups with different names in the Ethnographic Atlas and in the Census		
Name in Atlas	Name in Census	Source of matching
Aweikoma	Kaingang	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group mainly from Brazil. Amongst other ethnonyms Aweikoma is also known as Coroado, Cayapa, etc.
Aztec	Nahuatl	Encyclopedia Britannica. Available at http://www.britannica.com/ Amerindian group from Central Mexico. Nahuatl is the language spoken and disseminated by the Aztec Empire.
Caduveo	Kadiwéu, Guaikurú	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group from Paraguay and Brazil. Amongst other ethnonyms they are also known as Caduvi, Kaiwa, etc.
Campa	Ashaninka	Encyclopedia of World Cultures, Vol. VII –South America- and Ethnologue: Language of the World https://www.ethnologue.com/ Ashaninka belongs to one of the seven main groups of Amerindian group known as Campa. Ashaninkas tend to be found in remote areas in Peru and Brazil.
Cayapa	Chachi	Encyclopedia Britannica. Available at http://www.britannica.com/ Amerindian group from the west coasts of Ecuador that traditionally was known as Cayapas. Nowadays, they called themselves as Chachis.
Cayua	Guarani, Guarani Kaiowá, Guarani Mbya, Guarani Nhandeva, Ava-Guarani, Tupi-Guarani, Pal-Pavytera, Guarani Occident.	Métraux (1948). The Guarani. In Steward, Julian H. (ed.), Handbook of South American Indians, Vol. 3; and Ethnologue: Language of the World https://www.ethnologue.com/ Guarani encompasses a wider range of alternative names. Guarani can be found across Argentina, Bolivia, Paraguay and Brazil; and in each country the name of this group is used in different ways. For example, Bolivia classifies all type of Guarani-related groups under a single name known simply as Guarani – at least as it is reported in the recent census of 2012-. In Brazil, instead, it can be found three types of Guarani groups – Kaiowá, Mbya and Nhandeva-. However, all these names came to be labelled in the 18th century. Even the very single name of Guarani was introduced by the colonisers to differentiate those Amerindians that assimilated quickly the colonial power from those who rejected it, the latter in which case were identified by their pre-colonial name known as Cayua. A problem arises when in modern times those Amerindians who live in rural areas, and therefore have had hardly any external influence in their local matters, like to be called also as Guaranis. To overcome such a confusing association it seems that within the anthropological literature Guaranis are identified as Cayuas. And it is precisely this way that Murdock (1967) labelled them. We then proceed to aggregate all type of Guarani-related groups into a single classification: Cayuas.

CONTINUED (A.1)...

Panel B: Ethnic groups with different names in the Ethnographic Atlas and in the Census		
Name in Atlas	Name in Census	Source of matching
Chibcha	Muisca	Encyclopedia Britannica. Available at http://www.britannica.com/ Amerindian group mainly from Colombia. It is regarded as the second most influential group outside the Inca Empire in South America.
Choco	Embera	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group that lives in Colombia, Ecuador and Panama.
Cocopa	Cucapá	Amongst other ethnonyms they are also known as Cholo, Meme, Catio, etc.
Cuna	Tule	Ethnologue: Language of the World https://www.ethnologue.com/ Amerindian group from the northern Mexico who is also known as Kikima, Cocopah, etc.
Goajiro	Wayuu	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group that predominately lives in Panama but a few of them can also be found in Colombia. The name Cuna is more related to their cultural origins than the actual name they like to be used -Tule.
Guahibo	Sikuani	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group from Colombia and Venezuela. The name Goajiro is meant to have been introduced by Spanish colonisers.
Inca	Quechua	MEncyclopedia of World Cultures, Vol. VII –South America. Amerindian group from Colombia and Venezuela. The name Guahibo is from pre-colonial origins whereas Sikuani is a term labelled by them in recent times to encompass the whole groups under Guahibo heritage.
Jivaro	Achuar, Shuar, Shiwiar	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group that lives in Argentina, Chile, Bolivia, Ecuador and Peru. As its northern counterpart, Quechuas is the language spoken and disseminated by the Inca Empire. All these countries with Inca heritage identify this group as Quechuas.
Lengua	Enlhet, Enxet	Encyclopedia of World Cultures, Vol. VII –South America; and Ethnologue: Language of the World https://www.ethnologue.com/ Amerindian group mainly from Ecuador. They are also known as Givari, Zibaro, etc.
Paez	Nasa	Encyclopedia of World Cultures, Vol. VII –South America; and Ethnologue: Language of the World https://www.ethnologue.com/ One of the dozens hunter-gatherer bands of the Gran Chaco in Paraguay. The other ethnonyms used are Lengua-Sur and Lengua-Maskoi
Paraujano	Añu	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group that lives in Colombia. Paez is the language spoken by this group and instead they called themselves as Nasa.
Piapoco	Tzase	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group from Venezuela. Nowadays they called themselves as Añu. They are also known as Yapaco, Cuipoco, Deja, etc.

CONTINUED (A.1)...

Panel B: Ethnic groups with different names in the Ethnographic Atlas and in the Census		
Name in Atlas	Name in Census	Source of matching
Tarasco	Purépecha	Encyclopedia of World Cultures, Vol. VIII –Middle America and the Caribbean. Amerindian group from western Central Mexico. The term Purépecha comes from pre-colonial times whereas Tarasco was introduced during contact with Spanish colonisers.
Taulipang	Pemon	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group from Venezuela and Brazil. Amongst other ethnonyms they are also known as Arekuna or Taurepan.
Tucuna	Tikuna,	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group that live nearby the Amazon in Brazil, Colombia and Peru. Tucuna is a term with foreign origin. There are other ethnonyms to identify this group such as Jaunas, Tocunas, etc.
Tunebo	U'wa	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group mainly from Colombia. Amongst other ethnonyms they are also known as Covari, Luna, Tame, etc.
Warrua	Warao	Ethnologue: Language of the World https://www.ethnologue.com/ ; and Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group mainly from Venezuela. Warao is a self-name meaning "lowland people". Amongst other ethnonyms they are also known as Guarauno or Tuituna
Yahgan	Yámana	Ethnologue: Language of the World https://www.ethnologue.com/ Amerindian group that lives in Chile and Argentina. This group is also identified as Tequénica.
Yaruro	Pumé	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group from Venezuela. Pumé is a term commonly used amongst people within this group. On the other hand, Yaruro is meant to be a term mostly used by non-Amerindians in Venezuela. They are also known as Capuruchano and Saruro.
Yupa	Yukpa	Encyclopedia of World Cultures, Vol. VII –South America. Amerindian group from Venezuela. The difference between Yupa and Yukpa is only based on dialectic. This group is also identified as Yuko

Table A.2: Definitions and sources of variables (baseline analysis)

Variable	Description
<i>Dependent variables:</i>	
Infant Mortality Rates	The number of deaths of children under 1 year old in a given year per 1000 live births in each state. Source: Country's national statistics office
Drinking Water	Proportion of households in each state that have access to drinking water. Source: Redatam- Commission for Latin America and the Caribbean (ECLAC) and country's national statistics office
Electricity	Proportion of households in each state that have electricity. Source: Redatam- Commission for Latin America and the Caribbean (ECLAC) and country's national statistics office
Average Years of Schooling	Average of years of schooling from primary level onwards of population aged 15 and above in each state. Most recent value available in period between 1990 and 2006. Source: Gennaioli et al. (2013)
Primary Education	Proportion of total population in each state that completed primary education. Source: Redatam- Commission for Latin America and the Caribbean (ECLAC) and country's national statistics office
Secondary Education	Proportion of total population in each state that completed secondary education. Source: Redatam- Commission for Latin America and the Caribbean (ECLAC) and country's national statistics office
Log GDP per capita	Annual Log Gross Domestic Product per capita in each state as calculated by Bruhn and Gallegos (2012).
Poverty rates	Annual Log Poverty rates as calculated by Bruhn and Gallegos (2012).
<i>Main regressor of interest :</i>	
Index of Pre-colonial Institutions	As described in text. Source: Gray (1999) , A Corrected Ethnographic Atlas; Redatam- Commission for Latin America and the Caribbean (ECLAC) and country's national statistics offices
<i>Other independent variables:</i>	
Share of Amerindian Population	Proportion of Amerindian groups in total population in each state. Source: Redatam- Economic Commission for Latin America and the Caribbean (ECLAC) and country's national statistics offices
Population Density	Total population of each state divided by its total state's surface area (sq. km). Source: Redatam- Commission for Latin America and the Caribbean (ECLAC) and country's national statistics offices
Latitude	Absolute latitude of the centroid of each state. Source: Own calculation using Geographical Information System (GIS)
Altitude	Average altitude of each state (km). Source: Bruhn and Gallegos (2012) and Global Gazetteer Version 2.1

CONTINUED (A.2)...

Variable	Description
Temperature	Average temperature in degree Celsius in each state during period 1950-2000. Source: Gennaioli et al. (2013)
Land Suitability Index	Average land quality for agriculture in each state. Variable takes values between 0 and 1, with higher values denoting more fertile land. Data drawn from Ramankutty et al. (2002), who developed a geospatial index of land suitability for agriculture by examining three major components: croplands, climate conditions and soil characteristics. The index represents the probability that a particular grid cell will be cultivated. Ramankutty et al. (2002) report such index at a 0.5 degree resolution. We compute this index by spatially averaging all the grid cells that fall within each state. Data computed with ArcGIS 10.1.
Malaria Stability Index	Average values of malaria index. Variable takes values from 0 to 5, with higher values denoting more prevalence of the various types of malaria. Data drawn from Kiszewski et al (2004). To calculate this index we averaged all the grid cells that fall within each state. Data computed with ArcGIS 10.1
Land Area	Total surface area of each state in sq. km. Source: Country's national statistics offices
Land Locked	A dummy variable that indicates whether states have access to the sea. Source: Own calculation
Distance to Capital	Distance between the centroid of each state and the capital city of the country. Source: Own calculation
Inverse Distance to Coast	Distance between the centroid of each state and the nearest coastline in thousands of kilometres. Source: Gennaioli et al. (2013)
Oil & Gas	Dummy variable that indicates the existence of oil or gas fields in each state. Source: U.S. Geological Survey
Gold & Silver	Dummy variable that indicates the existence of gold or silver mines in each state. Source: U.S. Geological Survey
Other Mines	Dummy variable that indicates the existence of any other mines (copper, zinc, coal, etc.) in each state. Source: U.S. Geological Survey

Table A.3: Description of additional pre-colonial characteristics

Variable	Description
Gathering	Percentage of the population dedicated to the collection of wild plants and small land fauna. Source: Murdock (1967), Ethnographic Atlas, v1.
Hunting	Percentage of population dedicated to hunting. Source: Murdock (1967), Ethnographic Atlas, v2.
Fishing	Percentage of population dedicated to fishing. Source: Murdock (1967), Ethnographic Atlas, v3.
Agriculture	Percentage of population dedicated to agriculture. Source: Murdock (1967), Ethnographic Atlas, v5.
Settlement Pattern	An ordered variable that takes a value between 1-8, where 1 indicates fully nomadic and migratory societies; 2 is assigned to societies with semi-nomadic characteristics; 3 identifies societies with semi-sedentary characteristics; 4 indicates societies that lived in compact and impermanent settlements; 5 is given to societies those in neighbourhoods of dispersed family homes; 6 corresponds to groups in separated hamlets forming a single community; 7 is given to societies living in compact and relatively permanent settlements; and 8 corresponds to groups residing in complex settlements. Source: Murdock (1967), Ethnographic Atlas, v30.
Class Stratification	An ordered variable that takes a value between 1-5, where 1 is given to "absence of significant class distinction among freemen, ignoring variations in individual repute achieved through skill, valor, piety, or wisdom"; 2 corresponds to "wealth distinctions, based on the possession or distribution of property, present and socially important but not crystallized into distinct and hereditary social classes"; 3 is assigned to "elite stratification, in which an elite class derives its superior status from, and perpetuates it through, control over scarce resources, particularly land, and is thereby differentiated from a propertyless proletariat or serf class"; 4 indicates "dual stratification into a hereditary aristocracy and a lower class of ordinary commoners of freemen, where traditionally ascribed noble status is at least as decisive as control over scarce resources"; and 5 is given to "complex stratification into social classes correlated in large measure with extensive differentiation of occupational statuses". Source: Murdock (1967), Ethnographic Atlas, v67.
Slavery	A binary index that takes a value of 1 for societies characterized by any type of slavery and zero otherwise. Source: Murdock (1967), Ethnographic Atlas, v70.
Elections	A binary index that takes a value of 1 for societies where succession was conducted through "election or other formal consensus", and zero otherwise. Source: Murdock (1967), Ethnographic Atlas, v72.
Inheritance Rules for Property	A binary index that takes a value of 1 for societies reporting any type of inheritance rule of real property (land), and zero otherwise. Source: Murdock (1967), Ethnographic Atlas, v74

Table A.4: Contrasting rural and urban areas

Dependent variable	Primary education (1)	Secondary education (2)	Drinking water (3)	Electricity (4)
Pre-Colonial Institutions	-0.000495 [0.0111]	0.0149 [0.0355]	0.105** [0.0472]	0.112* [0.0628]
Pre-Colonial Institutions*Dummy_rural	0.0240** [0.0105]	0.0739*** [0.0210]	0.0491 [0.0339]	0.0561* [0.0314]
Controls included:				
Share of Amerindian pop.	YES	YES	YES	YES
Log population density	YES	YES	YES	YES
Controls for geography, location and natural resources	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES
Observations	581	581	581	581
Adjusted R-squared	0.828	0.784	0.548	0.565

Notes: Cluster standard errors at country level are in brackets. *** p<0.01, ** p<0.05, * p<0.1
Argentina has been removed from the sample for this table.

Appendix B

Appendix –*Chapter 3*–

Table B.1: Definitions and sources of variables

Variable	Description
<i>Main dependent variable:</i> Light Density at Night	Yearly average of light density during the period 2000-2013. This is a geospatial data that take discrete values ranging from 0 to 63, and is presented in raster files with a finer resolution of 30-second grid (similar as one squared kilometre). I compute this variable by averaging all the grids that fall within the ethnic homelands. Data computed on ArcGis 10.1. Source: Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS).
<i>Main regressor of interest:</i> Pre-Colonial Institutions	This variable measures the level of institutional advancements of ethnic groups on the basis of their homelands. To construct this variable I first had to digitise the traditional ethnic homelands in Latin America as of XVI. I then assigned the respective value of institutional complexity of the ethnic groups inhabited within these units. The variable takes discrete values between 0-2, where zero denotes groups with stateless features and 2 depicts groups with paramount chiefdom and state-related characteristics. Data computed in ArcGis 10.1 and Stata. Sources: <i>Outline of South American Cultures</i> (1957); <i>Ethnographic Bibliography of North America</i> (1975); and <i>Ethnographic Atlas</i> (1967)
<i>Controls:</i> Land Suitability Index	This is a geospatial index that examines the land suitability for agriculture through three major geographic components: croplands, climate, conditions and soil characteristics (Ramankutty et al. 2002). It takes values between 0 and 1, with higher values denoting more fertile land. I compute this index by averaging all the grids that fall within each ethnic homeland. Data computed in ArcGis 10.1 and Stata. Sources: .
Elevation	Data drawn from Ramankutty et al. (2002)
Temperature	By computing the average of all the grids in each ethnic homeland, I obtain the average elevation at the required level of analysis. Data computed in ArcGis 10.1 and Stata. Sources: http://www.sage.wisc.edu/atlas/data.php?incdataset=Topography
Precipitation	This is the mean yearly temperature data over the period 1950-2000. I average all the grids falling within my unit of analysis. Data computed in ArcGis 10.1 and Stata. Sources: Worldclim. Available at http://www.worldclim.org/current
Malaria	This is the mean yearly precipitation data (mm) over the period 1950-2000. I compute the data as my temperature variable above. Data computed in ArcGis 10.1 and Stata. Sources: Worldclim. Available at http://www.worldclim.org/current
Latitude	This is an index capturing the stability of malaria transmission across the world. I calculate the average of the grids for each ethnic homeland. Data computed in ArcGis 10.1 and Stata. Sources: Kiszewski A., et. al. 2004.
Oil & Gas:	I compute the absolute latitude in each ethnic homeland. Data computed in ArcGis 10.1.
	Using geospatial data drawn from the project called Peace Research Institute Oslo (PRIO), I create an indicator equals one if there is an oil or gas field within an ethnic homeland, and zero otherwise. Data computed in ArcGis 10.1 and Stata. Available at https://www.prio.org/

CONTINUED (B.1)...

Variable	Description
Population Density	Using the population digital data (as of 2000) from the United Nations Environment Program, I calculate the population density for each unit (in logs) by taking the average of all the grids as before. Data computed in ArcGis 10.1 and Stata. Available at http://na.unep.net/siouxfalls/datasets/datalist.php
Nearest Distance to Capital:	For each country, I calculate the nearest distance of ethnic homelands to the capitals. For example, using the centroid of each ethnic homeland within the Mexican territory, I compute their nearest distance (in kilometres) to Mexico City. Data computed in ArcGis 10.1 and Stata.
Pre-colonial population density	A variable taking the average of pre-colonial population density. I construct this variable by intersecting our map of ethnic homelands with a map containing all the contemporary boundaries of subnational states in Latin America. I then take the average of all the values of the subnational states falling in the ethnic homelands. Data computed on ArcGis 10.1 and Stata. Source: Bruhn and Gallegos (2012).
Share of Amerindian population	A variable taking the average of the present-day of share of Amerindian population. I construct this variable by intersecting our map of ethnic homelands with a map containing all the contemporary boundaries of subnational states in Latin America. I then take the average of all the values of the subnational states falling in the ethnic homelands. Data computed on ArcGis 10.1 and Stata. Source: own data.
Indicators on colonial economic activities	Four separate Indicators that take the average value of dummies falling in an ethnic homeland. These indicators report economic activity during the colonial period (other economic activities; none economic activities; mining; and plantation) as measured by Bruhn and Gallegos (2012). I construct these variables by intersecting our map of ethnic homelands with a map containing all the contemporary boundaries of subnational states in Latin America. I compute the average of the values of all the subnational states falling in an ethnic homelands.. Data computed on ArcGis 10.1 and Stata. Source: Bruhn and Gallegos (2012)

Appendix C

Appendix –*Chapter 4*–

Table C.1: Definitions and sources of variables

Variable	Description
<i>Main dependent outcomes</i>	
<i>Main regressor of interest:</i> Pre-Colonial Institutions	Our four outcomes of interest come from the 2010's Latinobarometer. The variables were constructed by using the question from the survey data asking the respondents how much trust they have in these institutions, for which they had to choose one of the following four ranking options: 1) <i>a lot</i> ; 2) <i>some</i> ; 3) <i>a little</i> ; and 4) <i>no trust</i> . We define a respondent who supports these institutions if he/she chose one of the first three options (<i>a lot</i> ; <i>some</i> or <i>a little</i>). Thus, an indicator is computed taking a value of one if individuals' responses are 1) <i>a lot</i> ; 2) <i>some</i> ; 3) <i>a little</i> ; and zero for the category 4) <i>no trust</i> . Data available at http://www.latinobarometro.org
<i>Individual-level controls</i>	This variable measures the level of institutional advancements of all the Amerindian groups living in each subnational state in Latin America. The index used two main sources: Ethnographic Atlas and subnational ethnic population statistics. The index was calculated according to the details on data and methodology outlined in chapter 1 of the present thesis. We assign to each respondent its respective value of the index of pre-colonial institutions using the location of subnational states as the main merging criteria.
<i>Subnational level controls</i>	All along the empirical analysis, a set of individual covariates were included as our present-day controls. These included age, age squared, gender, socioeconomic status, location and educational levels of respondents. All these variables were drawn from the 2010's Latinobarometer available at http://www.latinobarometro.org . Similarly, the empirical analysis included another set of demographic, geographical and historical controls, but all defined at subnational level. These variables are: Share of Amerindian population; Latitude; Altitude; Temperature; Land area (km. sq.); Distance to capital; Inverse distance to the sea; Landlocked; Oil and gas resources; Gold and silver resources; Population density (logs); Pre-colonial population density; Good colonial economic activities; Mining during colonial period; and Plantations during colonial period. All these variables come from the subnational dataset prepared for the first chapter of the present thesis. Definition of each variable can be consulted in previous Appendices.
<i>Dummies for each ethnic identity:</i> Asian Black Amerindian Mestizo Mulato White Others	This is a dummy variable taken a value of 1 if respondents identified themselves as having Asian origins, and zero otherwise. The information to compute such indicator was drawn from the 2010's Latinobarometer available at http://www.latinobarometro.org . This is a dummy variable taken a value of 1 if respondents identified themselves as having Black origins, and zero otherwise. The information to compute such indicator was drawn from the 2010's Latinobarometer available at http://www.latinobarometro.org . This is a dummy variable taken a value of 1 if respondents identified themselves as having Amerindian origins, and zero otherwise. The information to compute such indicator was drawn from the 2010's Latinobarometer available at http://www.latinobarometro.org . This is a dummy variable taken a value of 1 if respondents identified themselves as having Mestizo origins, and zero otherwise. The information to compute such indicator was drawn from the 2010's Latinobarometer available at http://www.latinobarometro.org . This is a dummy variable taken a value of 1 if respondents identified themselves as having Mulato origins, and zero otherwise. The information to compute such indicator was drawn from the 2010's Latinobarometer available at http://www.latinobarometro.org . This is a dummy variable taken a value of 1 if respondents identified themselves as having White origins, and zero otherwise. The information to compute such indicator was drawn from the 2010's Latinobarometer available at http://www.latinobarometro.org . This is a dummy variable taken a value of 1 if respondents identified themselves as having Other ethnic origins, and zero otherwise. The information to compute such indicator was drawn from the 2010's Latinobarometer available at http://www.latinobarometro.org